

A multi-locus perspective reveals connections between island biogeography and evolutionary
history of an endangered shrew (*Sorex pribilofensis*)

by

Benjamin J. Wiens

B.A., Bethel College, 2018

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Division of Biology
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2021

Approved by:

Major Professor
Andrew Hope

Copyright

© Benjamin Wiens 2021.

Abstract

Accelerating anthropogenic environmental change poses numerous threats to mammalian wildlife. Island endemics are among the most vulnerable species to rapid environmental change, and account for a disproportionate amount of all documented extinctions. The current vulnerability of island species to global change is a result of their unique evolutionary ecology. The evolutionary forces of natural selection, genetic drift, and reduced or non-existent gene flow lead to the high levels of endemism on islands, but can also leave these species vulnerable to change. It is therefore vital for island biodiversity conservation that we understand how past environmental change has influenced evolutionary dynamics. Islands in the Bering Sea represent a classic system of land-bridge insular evolution. Through Quaternary climate cycling, oscillating sea levels have alternately connected Alaska and Siberia through the Bering Isthmus, and today the Arctic is experiencing climate change at a more rapid pace than lower latitudes. St. Paul Island, home of the endemic Pribilof Island Shrew (*Sorex pribilofensis*), is located in the southern Bering Sea and has been isolated from the mainland for ~14,000 years. This shrew is part of a diverse sibling species-complex, which has a wide-ranging Holarctic distribution. The goals of my thesis are to (1) resolve the evolutionary relationship of *S. pribilofensis* to other related shrews, and (2) clarify the evolutionary processes leading to speciation among these enigmatic mammals.

Using a tiered genomic dataset of microsatellites, a maternally inherited mtDNA gene, and ~11,000 nuclear SNPs, I tested predictions related to the evolutionary and demographic history of *S. pribilofensis*. Given small island size and extended isolation, my overarching prediction was that genetic drift has led to rapid speciation and loss of genetic diversity within this shrew. In my first chapter, I show that *S. pribilofensis* is highly differentiated from sibling

taxa using Discriminate Analysis of Principle Components and Structure clustering analyses.

With phylogenetic analysis I then show that *S. pribilofensis* is the first to have diverged from closely related sibling taxa, and use Bayes Factor Species Delimitation to support species-level differentiation. In my second chapter, I give evidence for substantially reduced genetic diversity and a smaller effective population size of *S. pribilofensis* compared to mainland species.

Through multiple linear regressions, I then show that genetic differentiation is closely tied to reduced genetic diversity in this system of shrews. Finally, compared to mainland sibling shrews, *S. pribilofensis* is most strongly differentiated at its least variable loci and least differentiated at its most variable loci. These combined results are indicative of strong genetic drift driving the differentiation of *S. pribilofensis*. Put together, my findings indicate a scenario whereby *S. pribilofensis* speciated rapidly after island isolation through neutral divergence, but in the process has lost much of its genomic diversity. These results highlight the potential for genetic drift, as a consequence of dramatic environmental change, to rapidly reshape island biodiversity while as a result potentially leaving island species less able to respond to multiple additional environmental stressors in an anthropogenic world.

Table of Contents

List of Figures	vi
List of Tables	vii
Acknowledgements	viii
Chapter 1 - Introduction.....	1
Chapter 2 - Rapid allopatric divergence and speciation of an endangered insular shrew (<i>Sorex</i> <i>pribilofensis</i>)	5
Introduction.....	5
Methods.....	10
Results.....	19
Discussion	23
Figures.....	32
Tables.....	44
Chapter 3 - Genetic drift drives differentiation of an endangered insular shrew (<i>Sorex</i> <i>pribilofensis</i>)	49
Introduction.....	49
Methods.....	55
Results.....	63
Discussion	68
Figures.....	83
Tables.....	94
References.....	98
Appendix A - Rapid allopatric divergence and speciation of an endangered insular shrew (<i>Sorex</i> <i>pribilofensis</i>)	109
Supplementary Figures	109
Supplementary Table	110
Appendix B - Genetic drift drives differentiation of an endangered insular shrew (<i>Sorex</i> <i>pribilofensis</i>)	116
Supplementary Table	116

List of Figures

Figure 2.1 Distribution of shrews included in this study	32
Figure 2.2 ddRADseq parameter selection	33
Figure 2.3 Read depth per SNP locus	34
Figure 2.4 Microsatellite DAPC	35
Figure 2.5 SNP DAPC	36
Figure 2.6 Microsatellite Structure plots	37
Figure 2.7 SNP Structure plots	38
Figure 2.8 Cytb RAxML phylogeny	39
Figure 2.9 Cytb BEAST2 phylogeny	40
Figure 2.10 SNP RAxML phylogeny	41
Figure 2.11 SNP SVDquartets phylogeny	42
Figure 2.12 SNP Bayesian phylogeny	43
Figure 3.1 Detection of outliers with <i>PCadapt</i>	83
Figure 3.2 Detection of outliers with Bayescan	83
Figure 3.3 Neutral SNP DAPC	84
Figure 3.4 Non-neutral SNP DAPC	85
Figure 3.5 Neutral SNP Structure plots	86
Figure 3.6 Non-neutral SNP Structure plot	86
Figure 3.7 <i>S. pribilofensis</i> stairway plot	87
Figure 3.8 <i>S. cinereus</i> stairway plot	87
Figure 3.9 <i>S. portenkoi</i> stairway plot	88
Figure 3.10 <i>S. ugyunak</i> stairway plot	88
Figure 3.11 Combined stairway plots	89
Figure 3.12 Microsatellite linear regressions	90
Figure 3.13 Neutral SNP linear regressions	91
Figure 3.14 Polymorphic and monomorphic SNPs within <i>S. pribilofensis</i>	92
Figure 3.15 Pairwise genetic differentiation boxplots	93

List of Tables

Table 2.1 Shrew sample sizes	44
Table 2.2 ddRADseq missing data.....	44
Table 2.3 Species delimitation models	45
Table 2.4 Microsatellite genetic differentiation.....	46
Table 2.5 Cytb genetic differentiation	47
Table 2.6 SNP genetic differentiation.....	48
Table 2.7 Species delimitation.....	48
Table 3.1 Species of the <i>Sorex cinereus</i> complex.....	94
Table 3.2 Microsatellite effective population sizes	95
Table 3.3 Neutral SNP effective population sizes	95
Table 3.4 Genetic variability: microsatellites and cytochrome b	96
Table 3.5 Genetic variability: SNPs.....	96
Table 3.6 Genetic variability: neutral SNPs	97
Table 3.7 Genetic variability: polymorphic vs monomorphic SNPs	97

Acknowledgements

I would like to acknowledge all of the people and agencies involved in this research, whose many contributions made it possible. First, I thank my advisor, Andrew Hope, for all his support throughout my time at K-State. I have learned a lot from him during my two short years, and will be forever grateful for the opportunities he has given me. I also want to thank the members of my committee, Michi Tobler and Mark Ungerer. Their guidance and feedback have greatly contributed to my growth as a scientist.

This thesis would not have been possible without the innumerable contributions of those who came before me. For development and genotyping of the microsatellites used in this thesis, I thank Sandra L. Talbot, G. Kevin Sage, Sarah Sonsthagen, and Megan Gravley. I also thank the numerous field crews over the decades, in particular through the Beringian Coevolution Project, for the collection of shrews that were eventually included in this thesis, specifically, Joseph A. Cook and Eric P. Hoberg.

Additionally, I must thank Lauren Divine and Veronica Padula who, through their respective roles as Director of the Ecosystem Conservation Office and as Academic Program Director, Bering Sea Campus, greatly aided in my field work on St. Paul Island. They also provided the opportunity to connect with the Aleut community of St. Paul during the annual Bering Sea Days event. It was a wonderful experience, and I thank the native community for their welcoming attitude and great interest in my research. In particular, I thank the students at St. Paul school for their contributions to my work, from setting and checking traps to preparing shrew specimens for museum archival.

I also thank the agencies whose materials and funding made this thesis possible. Specimens and tissue samples included in this research were gratefully received on loan from

University of Alaska Museum of the North, Museum of Southwestern Biology, and Idaho Frozen Tissue Collection. Some of the computing for this project was performed on the Beocat Research Cluster at Kansas State University, which is funded in part by NSF grants CNS-1006860, EPS-1006860, EPS-0919443, ACI-1440548, CHE-1726332, and NIH P20GM113109. Additional financial support for my research was provided by the American Society of Mammalogists, the Central Plains Society of Mammalogists, and the Biology Graduate Student Association.

Finally, I thank my friends and family for all their support over the years. Despite my newfound interest in looking at dead mammals, they were always there for me. For them, I am forever grateful.

Chapter 1 - Introduction

The Bering Sea region has a history of dramatic environmental change through the Quaternary (the last three million years), and this continues today as the region experiences environmental change at a more rapid pace than lower latitudes (Bintanja et al., 2005; Stocker et al., 2013). These cyclical climate changes have been integral for shaping species' distributions, their history of diversification, and inter-specific interactions (Cook et al., 2016). In light of this, it is vital for the conservation of biodiversity in the region, particularly concerning species now associated with its isolated island systems, that we understand how past environmental change has influenced evolutionary dynamics (Colella et al., 2020). Environmental changes in the region surrounding the Bering Sea (Beringia) include oscillations in temperature and sea level, which have heavily influenced the landscape of Beringia as islands have alternately been connected and disconnected from the mainland (Bintanja et al., 2005). Populations historically isolated on these land-bridge islands face many natural stressors inherent to island isolation, including reduced range size, unique island climates, faunal relaxation and rapid community turnover, and extreme population cycling (Stuart et al., 2012; Weigelt et al., 2013). By interpreting genetic signatures, we can study how environmental trends have influenced species divergences, effective population sizes, and intra-specific genetic diversity, and infer continuing trends in response to modern environmental trajectories (Hope et al., 2015). From a functional evolutionary perspective, land-bridge island stressors often promote rapid evolutionary divergence from mainland populations by local adaptation through natural selection and/or by the random evolutionary process of genetic drift (Combe et al., 2021; Weigelt et al., 2016; Whittaker et al., 2008; Woolfit & Bromham, 2005). As anthropogenic environmental change intensifies, relict populations on these islands face the additional stressors of human

development, invasive species including potentially novel pathogens, and climate warming at an unprecedented rate (Leclerc et al., 2018; Russell & Kueffer, 2019; Stocker et al., 2013).

The Pribilof Island shrew (*Sorex pribilofensis*) represents a classic example of land-bridge island evolution. Endemic to St. Paul Island in the Bering Sea and last isolated from mainland Alaska ~14 kya, this shrew faces many of the natural and anthropogenic stressors associated with environmental change in the region. It has experienced a rapid reduction in range size as St. Paul Island shrunk to ~71 km² since initial isolation, faunal relaxation and loss of multiple community interactions, recent human development and introduction of non-native species, possible novel pathogen interactions, and a rapidly warming island climate (Combe et al., 2021; Graham et al., 2016; Hope et al., 2016; Mason 2018). Currently listed as an endangered species, it is part of a species complex of closely related shrews that occur throughout Beringia. The phylogeographic history of these shrews is closely linked to the glacial-interglacial cycles of the Quaternary; most recently, since the last glacial maximum (LGM) ~20 kya, inundation of the Bering Isthmus has resulted in range fragmentation of an ancestral shrew taxon followed by allopatric divergence through geographic isolation on islands and on separate continents, resulting in the current distribution of multiple sister taxa (Hope et al., 2012). As such, *S. pribilofensis* in the context of this system provides an excellent opportunity for which to investigate the evolutionary consequences of insular endemism and what appears to be a case of remarkably rapid speciation among mammals.

My thesis takes advantage of the opportunity that these focal shrews provide for investigating multiple aspects of the evolution of Beringian wildlife, and I present two overarching research goals. In my first research chapter, I aim to estimate the degree of genetic differentiation within this system, clarify phylogenetic systematic relationships among sibling

shrew taxa, and explicitly test the species designation of *S. pribilofensis*. Given that *S. pribilofensis* is federally listed as endangered, a detailed understanding of taxonomic relationships and levels of genetic divergence is crucial for ongoing management. I predict that *S. pribilofensis* will exhibit high genetic differentiation, be phylogenetically distinct, and that its designation as a distinct species will be supported. In my second research chapter, I aim to increase our understanding of the evolutionary forces that lead to the unique biodiversity of land-bridge islands. Considering that all land-bridge island endemics in the Bering Sea must respond to accelerating environmental change, knowledge of the evolutionary forces driving the differentiation of this shrew can be used to inform predictions for both species-specific and general evolutionary trajectories of biodiversity in Beringia. Due to rapid geographic isolation on a shrinking island, I predict that genetic drift has been the predominant driving force in the evolution of *S. pribilofensis*.

To address these goals, I make use of three independent genetic datasets, including maternally inherited mitochondrial cytochrome b sequences (1,140 bp), 20 microsatellite loci, and ~11,000 single nucleotide polymorphisms (SNPs). I test predictions related to the evolutionary and demographic history of *S. pribilofensis* under a coupled phylogenetic and population genetic framework. My findings support the recognition of *S. pribilofensis* as a distinct species and provide evidence for genetic drift driving speciation in this system. Put together, these results indicate a scenario whereby *S. pribilofensis* diverged rapidly after island isolation through neutral processes, but this was accompanied by a substantial loss of genomic diversity. Given the evolutionarily recent geographic isolation of *S. pribilofensis*, these combined results indicate the potential for strong genetic drift to rapidly reshape the biodiversity of land-bridge islands. My results further our understanding of how species on Arctic islands respond to

dramatic environmental change, and show that this response can leave them vulnerable to continued and accelerating anthropogenic changes. This highlights *S. pribilofensis* as a species of high conservation concern with little remaining genetic variability through which to adapt to new environments, in a rapidly changing world.

Chapter 2 - Rapid allopatric divergence and speciation of an endangered insular shrew (*Sorex pribilofensis*)

Introduction

Earth's biodiversity is currently facing a sixth mass extinction (Ceballos et al., 2015). Central to conservation of biodiversity in the face of this global phenomenon is building an underlying understanding of systematic relationships and nomenclature that effectively describes significant evolutionary lineages (Soltis & Gitzendanner, 1999; Ely et al., 2017). Inaccurate species designations can lead to management practices that focus on the wrong level of diversity, to the detriment of the focal taxa (Ely et al., 2017). Conservation efforts for an under-split species (i.e. a taxon comprised of multiple unrecognized species) may not have the resolution to account for the true biological units that warrant conservation (Soltis & Gitzendanner, 1999). Successful identification of cryptic but significant evolutionary lineages has become increasingly more common with the advent of molecular genetics and next-generation sequencing (Fišer et al., 2018; Graham et al., 2017). These methods in combination with the unified species concept, which describes all independently evolving lineages as species, and uses tenets of other species concepts as lines of evidence for identifying such lineages, allows for accurate classification and conservation of the world's biodiversity (de Queiroz, 2005; Soltis & Gitzendanner, 1999). In addition to informing classification at the species level, phylogenetic and population genetic analyses provide further insight into lineage and population dynamics, which can be integral focal units for maintaining species diversity and which may warrant targeted conservation efforts of their own (Soltis & Gitzendanner, 1999).

Islands account for a disproportionately large amount of the world's biodiversity, and island species are disproportionally threatened by anthropogenic environmental change (Tershy

et al., 2015). Through founder effects, population bottlenecks, and subsequent adaptation to novel environments, island species may experience rapid evolution, which can lead to often rapid divergence, endemism, and speciation (MacArthur & Wilson, 1967). Stochastic environmental events, particularly on small islands, can increase the likelihood of sudden and dramatic population fluctuations. Relatively small baseline population sizes in addition to these fluctuations can lead to the progressive loss of genetic diversity through strong genetic drift (Nei et al., 1975). A continued reduction in genetic diversity eventually leads to limited variability for selection to act upon, resulting in a decrease in adaptive potential (Agashe et al., 2011; Reed & Frankham, 2003). Neutral evolutionary processes are often accompanied to varying degrees by natural selection, leading to endemism through time in response to the local environments of the islands on which they occur. The persistence and conservation of island species in the face of modern environmental change will therefore depend on their unique evolutionary histories and on accurate species designation of rapidly evolving lineages.

No matter whether forces of drift or selection predominate, island biodiversity that has evolved under a specific range of environmental conditions is often highly susceptible to rapid, sudden, or extreme changes to island environments. Thus, the same forces that lead to the evolution of distinct island species also lead to their susceptibility to rapid environmental change, variably reflecting reduced genetic diversity and/or adaptation to specific conditions (Reed & Frankham, 2003; Weigelt et al., 2013). The accelerating pace of global anthropogenic environmental changes may alter local conditions beyond the limitations of island species. It is therefore vital to investigate island endemics through development of integrated research and management that works to recognize distinct island biodiversity and accounts for the unique evolutionary forces that shape it (Graham et al., 2017; Kueffer et al., 2014).

The Pribilof Island shrew (*Sorex pribilofensis*) is an endangered species endemic to St. Paul Island, located in the Bering Sea. This region has a history of dramatic environmental change since the Quaternary; sea level oscillations due to glacial-interglacial cycling have alternately connected Alaska and Siberia through the Bering Isthmus during extended cold phases and lowered sea levels, and then disconnected the continents during more brief warm phases with higher sea levels leaving only island remnants (Bintanja et al., 2005; Elias & Brigham-Grette, 2013; Jakobsson et al., 2017). St. Paul Island, part of the Pribilof Islands group, is one such land-bridge island. As sea levels began to rise following the last glacial maximum (LGM) ~20 kya, it was one of the first land masses to be isolated from the Bering Isthmus, disconnecting ~14 kya (Graham et al., 2016). At this time, St. Paul likely supported a broad subset of Beringian species. Following isolation the island initially decreased in size rapidly, until about 9 kya. The island then continued to shrink at a slower pace until about 6 kya at which point it reached its present size of only 71 km² (Byrd & Norvell, 1993; Graham et al., 2016). Through this period, diversity on the island would have relaxed through progressive extinction of species as space, resources, and vital species connections were reduced or lost (Combe et al., 2021). It is possible that the relaxation phase of St. Paul biodiversity continues, with further species extinctions a logical theoretical expectation within this system. In addition to its small size, St. Paul Island is highly remote, being isolated from mainland Alaska by ~450 km and from the Aleutian Islands by >400 km. Today, the environment on St. Paul is characterized by sub-Arctic climate, maritime tundra vegetation, the absence of trees and shrubs, no permafrost, a few freshwater lakes, no springs or streams, and low relief (max elevation 203 meters above sea level) (Byrd & Norvell, 1993; Colinvaux, 1981).

Initial scientific knowledge of *S. pribilofensis* came from three expeditions to St. Paul island in the 1900s, which all found it to be abundant on the island (Byrd & Norvell, 1993; Fay & Sease, 1985; Jackson, 1928). Byrd & Norvell (1993) found the shrew to be most abundant in tall-plant vegetation, specifically in dune and grass-umbel habitats, and did not capture any shrews in wet sedge or short-plant habitats. They calculated that the preferred habitat of the shrew covers 54.9% of the island, or 39.0 km². Morphological divergence from related mainland shrews earned *S. pribilofensis* its specific status (Yudin, 1969; Merriam, 1895; van Zyll de Jong, 1982), and subsequent preliminary analysis of one mitochondrial and a few nuclear genes have shown it to be monophyletic, though its relationship to other Beringian shrews remains poorly resolved (Demboski & Cook, 2003; Hope et al., 2012). *Sorex pribilofensis* is a part of the *Sorex cinereus* species complex (hereafter *cinereus* complex), which currently consists of 13 species of the subgenus *Otisorex* (Hutterer, 2005), broken up into two reciprocally monophyletic clades (Beringian and Southern; van Zyll de Jong, 1991; Demboski & Cook, 2003). The Beringian clade (the focus of my research) currently contains 9 species, though these species exhibit very shallow divergence based on mitochondrial and nuclear genetic markers, making their specific status questionable (Hope et al., 2012). Of these 9 species, 6 occur in the Beringian region: *Sorex ugyunak* in northernmost Alaska and Canada; *Sorex jacksoni*, restricted to St. Lawrence Island; *Sorex pribilofensis*, restricted to St. Paul Island; *Sorex portenkoi* and *Sorex camtschatica* in easternmost Russia; and *Sorex leucogaster*, restricted to Paramushir Island, south of the Kamchatka Peninsula. The three remaining species, *Sorex haydeni*, *Sorex preblei*, and *Sorex lyelli* occur outside of Beringia (northern/central North American prairies, the Columbia basin, and Sierra Nevada, respectively) but are still considered within the Beringian clade based on their evolutionary relationship to northern taxa. The phylogeography of the six northern species

distributed through Beringia is consistent with a history of fragmentation of a single ancestral taxon into multiple isolated areas, dating to the break-up of the Bering Isthmus, followed by allopatric divergence (Hope et al., 2012). As fragmentation of this ancestral taxon occurred very recently from an evolutionary perspective (only 5-14 kya), it is surprising that these taxa are afforded status as distinct species, because rates of speciation for mammals are normally orders of magnitude slower (Barnosky, 2005).

Today, the Bering region is experiencing climate change at a more rapid pace than lower latitudes (Stocker et al., 2013). These changes include a rapid increase of mean temperature, reduction of sea-ice extent, rising sea levels, and the advance of boreal-associated temperate biodiversity encroaching on sub-Arctic and Arctic tundra habitats (Callaghan et al., 2004; Stocker et al., 2013). This climate-driven community turnover presents a threat to the endemic biodiversity of Beringia, as species must either adapt to changing environments, shift their range to track suitable environmental conditions (not a viable option for most insular terrestrial species), or face extinction (Colella et al., 2020). Conservation efforts that take into account the current evolutionary trajectories of these species can contribute to their persistence. However, in order for conservation to be effective, we must have a detailed understanding of the taxonomy and species relationships of the focal taxa.

In this chapter I test the specific status of *S. pribilofensis* under a population genetic and phylogenetic framework, in order to better inform current conservation status for this data-deficient species (IUCN, 2016). To this end, I make use of three independent genetic datasets consisting of nuclear microsatellites, mitochondrial cytochrome b sequences, and ~11,000 nuclear single nucleotide polymorphisms (SNPs), to investigate species relationships within the *cinereus* complex at higher resolution than previous studies. First, I test the level of genetic

differentiation between these taxa, characterize species relationships through analysis of genetic structure through this system, and use multiple methods to infer phylogenetic relationships in the *cinereus* complex. Based on island biogeographic theory, I predict (1) strong genetic differentiation of *S. pribilofensis* from sister taxa, reflecting longer and more extreme (small-island) isolation, (2) independent clustering of each Beringian shrew taxon based on structure analyses, indicating independent evolutionary trajectories through allopatry, and (3) clear monophyly and specific status of *Sorex pribilofensis* relative to the rest of the Beringian clade. I then use Bayes Factor species Delimitation (BFD) to compare multiple alternate species delimitation models in order to explicitly test the species designation of *Sorex pribilofensis*.

Methods

Sample collection and DNA extraction

Shrews were collected from throughout Beringia between the years 1952-2019, using both pitfall and “museum special” snap traps, and all specimens were fully processed and archived in publicly accessible museum collections according to published guidelines (Galbreath et al., 2019). For the purposes of this study, a total of five species of the Beringian clade of the *cinereus* complex were included: *Sorex pribilofensis*, *Sorex ugyunak*, *Sorex portenkoi*, *Sorex camtschatica*, and *Sorex jacksoni*. No samples of *S. leucogaster* were available for genetic analyses. Additionally, *Sorex cinereus* (*sensu stricto*) was included to serve as an outgroup. Sample sizes of each species used to generate three genetic datasets are similar and mostly overlapping (Table 2.1). Specimen museum numbers and localities are provided in a Supplementary Table (Table A1). The distributions of these shrews collectively occupy most of Beringia, although their ranges are allopatric except for a zone of sympatry in northern Alaska between *S. cinereus* and *S. ugyunak* (Fig. 2.1). Shrew tissue samples were preserved frozen at -

80°C or in 95% ethanol. DNA was extracted from liver or muscle samples using slightly modified standard salt extraction methods (Miller et al., 1988).

Microsatellite PCR, genotyping, and data filtering

For representatives of all 6 species, 20 microsatellite loci (developed and characterized by Sonsthagen et al., 2013) were amplified through polymerase chain reaction (PCR) following the amplification and thermocycler conditions of Toussaint et al. (2012). Scoring errors and potential null alleles were identified with Micro-checker (Van Oosterhout et al., 2004). Prior to analyses, individuals with a high proportion (>50%) of missing data were removed from the dataset. Then, loci for which all individuals of any given species had missing data, as well as loci with >20% missing data across all species, were removed. The remaining loci were then independently tested for deviation from Hardy-Weinberg equilibrium, separately in each species, using the exact test and 10,000 Monte Carlo permutations implemented in the R package *pegas* (Guo & Thompson, 1992; Paradis, 2010). Any loci that deviated significantly ($p < 0.05$) in >3 species were removed. The final dataset included 17 loci for 79 individuals across 6 species (Table A1).

Cytochrome b PCR, sequencing, and data filtering

The mitochondrial cytochrome b (Cytb) gene was amplified through PCR using previously optimized primers (MSB05, MSB14) and thermocycling conditions (Hope et al., 2010). Sanger sequencing was performed on an ABI 3730 by Genewiz LLC (South Plainfield, NJ). Raw reads were cleaned and aligned using Geneious (Kearse et al., 2012). All previously sequenced and publicly available Cytb sequences for *S. camtschatica* and *S. jacksoni* were downloaded from GenBank and included in this dataset. For standardizing analyses across

genetic datasets, publicly available *Cytb* sequences for *S. cinereus*, *S. pribilofensis*, *S. portenkoi*, and *S. ugyunak* individuals were only included in the *Cytb* dataset if that individual was represented in one or both of the microsatellite and SNP datasets. Prior to analyses, individuals with >25% missing data (i.e. >285 missing base pairs out of the full 1140 bp sequence) were removed. The final *Cytb* dataset included sequences for 104 individuals across 6 species (Table A1).

ddRADseq library prep and sequencing

Genomic DNA was quantified using Quant-iT Picogreen dsDNA Assay (Invitrogen) and gel electrophoresis (2% agarose) to identify 91 samples with sufficient yields (>100ng) of high molecular weight DNA, which were then submitted to the University of Minnesota Genomics Center (UMGC), Minneapolis for double digest restriction site associated DNA sequencing (ddRADseq; Peterson et al., 2012). UMGc prepared ddRADseq libraries and sequenced samples using the following protocols. For each sample, 100 ng of DNA was digested with 10 units each of *SbfI* and *TaqI* restriction enzymes from New England Biolabs (NEB) and incubated at 37 °C for 2 hours before heat inactivating at 80 °C for 20 minutes. Samples were then ligated with 200 units of T4 ligase (NEB) and phased adaptors with CRYG and CG overhangs at 22 °C for 1 hour before heat killing. The ligated samples were purified with SPRI beads and then amplified for 18 cycles with 2x NEB Taq Master Mix to add unique barcodes to each sample. Libraries were purified, quantified, pooled, and size selected for the 300 – 744 bp library region and diluted to 2 nM prior to sequencing. UMGc sequenced 150-bp single-end reads across 0.25 lanes of a NextSeq 550 High-Output FlowCell (Illumina, USA). The resulting fastq files were demultiplexed using Illumina bcl2fastq software and Trimmomatic (Bolger et al., 2014) and Cutadapt (Martin, 2011) were used to remove the padding sequence at the beginning of the read,

trim the Illumina adapter of the 3' end, and discard reads shorter than 93bp. The trimmed reads were then filtered for quality using the *process_radtags* module in Stacks v2.54 (Rochette et al., 2019). Reads with uncalled bases were removed, and a sliding window approach was used to drop reads with low quality scores (phred score <10). Following these steps, 37,442,922 total reads were retained.

Denovo RAD locus and SNP identification

Loci were discovered *de novo* using the *denovo_map.pl* pipeline in Stacks v2.54. First, the parameters controlling loci formation and polymorphism were optimized for the dataset following the recommendations provided by the software developers (Rochette & Catchen, 2017), and a locus catalog was built using the optimal parameters (Fig. 2.2). Single nucleotide polymorphisms (SNPs) were then called for each locus. A total of 89,179 polymorphic SNPs were identified from 59,328 loci, with an average read depth of 14.0. These loci and SNPs subsequently went through several quality filters. Loci had to be present in >80% of individuals within a species to be processed for that species. Loci then had to be present in at least 3 species to be processed for the entire dataset. For loci with multiple SNPs, only the first was retained to minimize the potential for linkage. SNPs with a minor allele count <2 were removed, as these may reflect PCR error. SNPs with an observed heterozygosity >70% were also removed as they may represent paralogous loci. The read depth and proportion of missing data per individual was then assessed in the R packages *vcfR* (Knaus & Grünwald, 2017) and *SambaR* (Jong et al., 2021) (Fig. 2.3; Table 2.2). Individuals with >25% missing data were removed from the dataset. For each species, each SNP was then tested separately for deviation from Hardy-Weinberg equilibrium using the exact test and 100 Monte Carlo permutations implemented in the R package *pegas* (Guo & Thompson, 1992; Paradis, 2010). SNPs that were out of Hardy-

Weinberg equilibrium in >2 species were removed. These filters retained a total of 11,989 polymorphic SNPs for 90 individuals across 4 species (Table A1).

Genetic differentiation statistics

Genetic differentiation of Beringian shrew taxa was assessed using multiple statistics for each of the microsatellite, Cytb, and SNP datasets. For each taxon pair in the microsatellite and SNP datasets, I calculated pairwise F_{ST} , G_{ST} , and Jost's D values using the R packages *hierfstat* (Goudet, 2005) and *mmod* (Winter, 2012). Specifically, F_{ST} was calculated in *hierfstat* following Nei (1987). For the Cytb dataset I estimated pairwise F_{ST} values and percent haplotype divergences between each taxon pair in DnaSP v6.12 (Rozas et al., 2017).

Assignment of independent evolutionary clusters

Both the microsatellite and SNP datasets were used to analyze species structure through two clustering methods. First, Discriminate Analysis of Principle Components (DAPC) was implemented in the R package *adeigenet* (Jombart, 2008). Groups were assigned *a priori* based on recognized current taxonomy. For both datasets, an initial DAPC was run retaining the number of principle components (PCs) that explained ~90% of the cumulative variance and retaining all of the discriminant functions. Using this initial DAPC, the optimal number of PCs to retain was then determined using the *a.optim.score* function with 10,000 simulations for each possible number of retained PCs. A final DAPC was then run for both datasets with the optimal number of PCs retained and all discriminant functions retained.

Species structure was also analyzed through a model-based clustering method implemented in Structure v2.3.4 (Pritchard et al., 2000). For the microsatellite dataset, I tested scenarios for a range of numbers of genetic clusters ($K = 2-8$) and ran 10 repetitions for each

value of K. I used a burn-in length of 500,000 Markov chain Monte Carlo (MCMC) iterations, followed by a run length of 1,000,000 iterations. I used the correlated allele frequency model and the admixture model, and performed these runs both with and without current species designations as a set prior to compare the results. For the SNP dataset, I tested scenarios for K=2-5 and ran 10 repetitions for each value of K. The structure runs were multithreaded using the software EasyParallel (Zhao et al., 2020). I used a burn-in length of 100,000 MCMC iterations followed by a run length of 500,000 iterations, and again performed these runs both with and without current species designations as a set prior. For all runs, the most likely number of genetic clusters was selected using the Evanno method (Evanno et al., 2005) implemented in the software Structure Harvester (Earl & vonHoldt, 2012). All structure plots were visualized using the software Structure Plot (Ramasamy et al., 2014).

Phylogenetic analyses: mtDNA Cytb gene

Phylogenetic trees were inferred for the full Cytb dataset using two separate methods. I first used a maximum likelihood (ML) method to estimate the phylogeny using RAxML v8.2.12 (Stamatakis, 2014), implementing the GTR+CAT nucleotide substitution model with 1,000 bootstrap replicates. The resulting tree was rooted at the midpoint and visualized in FigTree v1.4.4 (Rambaut & Drummond, 2012). In the second method, I inferred the Bayesian multispecies coalescent tree using BEAST v2.6.3 (Bouckaert et al., 2019). I inferred the nucleotide substitution model, range of rate heterogeneity, and proportion of invariant sites simultaneously during the MCMC analysis with the bModelTest package (Bouckaert & Drummond, 2017), with the transition-transversion split option and empirical frequencies. I used a lognormal relaxed molecular clock with a mutation rate of 5.5% per million years (Hope et al., 2010) and the Coalescent Constant Population tree prior. I ran the MCMC chain for 10,000,000

generations sampling every 5,000 generations and used Tracer v1.7.1 (Rambaut, 2018) to confirm the convergence of the chain (i.e. ESS values for parameters >300). I used TreeAnnotator v2.6.2 (Rambaut & Drummond, 2020) to drop the first 10% of trees as burn-in and to annotate the consensus tree, then visualized the tree topology in FigTree v1.4.4.

Phylogenetic analyses: SNPs

Phylogenetic trees were inferred for the SNP dataset using three separate methods. The first considered all individuals independently but analyzed all SNPs as a single concatenated dataset; the second generated a species tree considering all individuals and all SNPs independently; the third estimated a species tree considering individuals assigned to *a priori* “species” designations, but considering independent genealogies for each SNP.

For the first method I estimated the ML phylogeny using RAxML v8.2.12. To estimate the ML tree, the SNPs were first converted to a phylip format, using ambiguity codes for heterozygous sites and Ns for uncalled sites following standard IUPAC notation. An ascertainment bias correction was applied to the likelihood calculation, as is recommended when constructing phylogenies with SNPs, due to the lack of invariant sites (Leaché et al., 2015). However, RAxML considers sites to be invariant if phasing the IUPAC characters could lead to invariant sites; therefore I filtered out these potentially invariant sites using the python script (https://github.com/btmartin721/raxml_ascbias), yielding 9,144 variant sites. I then estimated the ML phylogeny with 100 bootstrap replicates, implementing the GTR+G nucleotide substitution model and the Lewis method for ascertainment bias correction. The resulting tree was rooted at the midpoint and visualized in FigTree v1.4.4.

In the second method, I inferred a species tree in SVDquartets (Chifman & Kubatko, 2014) implemented in PAUP v4.0 (Swofford, 2017). This method infers the relationships among

quartets of taxa under the coalescent model before assembling the quartets into a species tree. I randomly sampled 1,000,000 quartets (representing 39.14% of the total possible quartets) and measured the uncertainty with 200 non-parametric bootstrap replicates. The tree was rooted at the midpoint and tree branch lengths were optimized using maximum likelihood under the GTR+I+G model. The tree topology was visualized in FigTree v1.4.4.

In the third method, I inferred the Bayesian coalescent species tree with SNAPP v1.5 (Bryant et al., 2012) implemented in Beast v2.6.3. Due to computational restraints, only 8 individuals per species were selected for analysis, and SNPs were filtered to exclude invariant sites based on a reduced sample size. This resulted in 11,387 polymorphic sites retained across 32 individuals. Due to ascertainment bias, it is difficult to determine absolute mutation rates for a SNP dataset, so mutation rates U (forward mutation rate) and V (backward mutation rate) were fixed at 1.0. The coalescence rate was initially set at 1.0 but was sampled throughout the analysis and the rate prior distribution was uniform. I ran the MCMC chain for 10,000,000 generations sampling every 1,000 generations and used Tracer 1.7.1 to confirm the convergence of the chain (i.e. ESS values for parameters >300). I used TreeAnnotator 2.6.2 to drop the first 50% of trees as a burn-in and to annotate the consensus tree, then visualized the tree topology in FigTree v1.4.4.

Species delimitation

To test support for the species designation of *S. pribilofensis*, I statistically compared three alternate species delimitation models using the Bayes Factor species Delimitation (BFD) method (Leaché et al., 2014), implemented by SNAPP v1.5 in BEAST v2.6.3. This method allows for comparisons of alternate species delimitation models under an explicit multispecies coalescent framework, making use of genome-wide SNP data. Based on the observed species

structure from DAPC and Structure analyses, as well as the observed phylogenetic relationships based on the Cytb and SNP datasets, I tested three species delimitation models using the SNP dataset (Table 2.3). In the first model four species were assigned, with each currently recognized taxon treated as independent species. In the second model, reflecting the results of structure clustering analyses, three species were assigned, with *S. ugyunak* and *S. portenkoi* treated as one species, and *S. pribilofensis* and *S. cinereus* each treated as distinct. In the third model only two species were assigned, with *S. cinereus* recognized as one species and the remaining three treated as one species.

Due to computational restraints, I only used 16 (4 per taxon) of the 32 individuals included in the estimation of the initial SNAPP phylogeny, and SNPs were again filtered to exclude invariant sites based on a reduced sample size. This resulted in 10,368 polymorphic sites retained across 16 individuals. In order to provide proper priors on the speciation rate (λ) and theta (θ) for the BFD analyses, I first estimated a time-calibrated species tree following the divergence-time estimation method outlined by Stange et al. (2018). This model employs a strict molecular clock, links all ancestral and contemporary populations sizes, and assumes a Jukes-Cantor model of nucleotide substitution. To time-calibrate the molecular clock, I interpreted the formation of St. Paul Island (~14 kya; Graham et al., 2016) and the formation of the Bering Strait (~11 kya; Jakobsson et al., 2017) as the vicariance events causing first the divergence of *S. pribilofensis* (St. Paul Island), followed by the divergence of *S. ugyunak* (Alaska) from *S. portenkoi* (Siberia). Accordingly, I set a uniform prior distribution of 13-15 kya on the node age of the *S. pribilofensis*/*S. ugyunak*/*S. portenkoi* clade and a uniform prior distribution of 10-12 kya on the node age of the *S. ugyunak*/*S. portenkoi* clade. The topology of the original, undated SNAPP species tree was used to create a starting tree with node ages of the two calibration nodes

set at 14 kya and 11 kya. Input files for SNAPP were prepared with the script `snapp_prep.rb` (https://github.com/mmatschiner/snapp_prep). This script removed sites for which there was only missing data in one or more species, which left a total of 7,801 sites for analysis. I ran the MCMC chain for 15,000,000 generations sampling every 5,000 generations and removing a 10% burn-in. The population size estimates were added to the log file with the script `add_theta_to_log.rb` (https://github.com/mmatschiner/snapp_prep) and then I used Tracer 1.7.1 to confirm the convergence of the chain and the convergence of the λ and θ parameters (i.e. ESS>300).

After initial estimation of λ and θ , I then calculated the marginal likelihood estimates (MLE) of each species delimitation model, with the individuals grouped into species accordingly for each species model (Table 2.3). The mutation rates U and V were fixed at 1.0 and the coalescence rate was initially set at 1.0 but was sampled throughout the analysis. I used gamma prior distributions for both λ and θ with the mean values for each set at the estimated means calibrated during the divergence-time estimation analysis. I ran the MCMC chain for 100,000 generations sampling every 1,000 generations. I conducted 48 path sampling steps, each with 50,000 MCMC steps, a pre-burnin of 5,000, and alpha set at 0.3, using the Model_Selection v1.5.3 package in BEAST v2.6.3, then confirmed the convergence of the chain (i.e. ESS>300) for each path sampling step. I ranked the alternative species delimitation models according to their MLE and calculated Bayes factors ($BF = 2 \times (MLE_1 - MLE_2)$) to compare the models (Leaché & Bouckaert, 2018).

Results

Genetic differentiation

I found varying levels of genetic divergence among taxa, although relative levels of differentiation were generally consistent across all four measures (F_{ST} , G_{ST} , Jost's D , and percent haplotype divergence; Tables 2.4-2.6). For all three genetic datasets, *S. cinereus* was consistently most divergent from other taxa, closely followed by *S. pribilofensis*. Based on all measures, *S. ugyunak* and *S. portenkoi* were least genetically differentiated of any taxon pair. For *S. pribilofensis*, pairwise genetic differentiation values were highest in relation to *S. cinereus*, except for F_{ST} and G_{ST} values based on the microsatellite dataset, which were highest in relation to *S. jacksoni*. Based on all genetic datasets and measures of genetic differentiation, *S. pribilofensis* was most genetically similar to *S. ugyunak*.

Independent evolutionary clusters

For Discriminate Analysis of Principle Components (DAPC), the optimal number of principle components (PCs) to retain for the microsatellite and SNP datasets were 7 and 3, respectively (Figs. 2.4, 2.5). When plotting all taxa using the first two PCs, both *S. pribilofensis* and *S. cinereus* were independently isolated from other taxa, and *S. cinereus* was the most distant group relative to the other taxa, followed by *S. pribilofensis*. However, based on the microsatellites, *S. cinereus* was only slightly more distant than *S. pribilofensis*. For both datasets, *S. portenkoi* and *S. ugyunak* group very closely to each other, and based on microsatellite data are also closely aligned with *S. jacksoni* and *S. camtschatica*. While *S. portenkoi* and *S. ugyunak* group closely for both genetic markers, there is more distinction between the two using the SNP dataset. However, while there is not much distinction between *S. portenkoi*, *S. ugyunak*, *S. camtschatica*, and *S. jacksoni* when using the first two PCs, the following PCs (3 and 4 for the microsatellite dataset, 3 for the SNP dataset) better distinguish these taxa from each other, while losing power to distinguish between *S. cinereus* and *S. pribilofensis*.

The relationships based on DAPC analyses are also evident from clustering analyses using Structure (Figs. 2.6, 2.7). For both microsatellite and SNP datasets, using *a priori* species designations did not impact how individuals clustered or the total number of clusters, so here I report results for Structure analyses without *a priori* species designations. Based on the Evanno method, the best supported value of K for both the microsatellite and SNP data was K=3, although for both datasets K=5 also had a high likelihood (Figs. A1, A2). However, under K=5 for the SNP dataset, no measurable portion of any individual was assigned to the fifth cluster, so K=5 essentially reflected 4 clusters. For both datasets, when K=3 *S. pribilofensis* and *S. cinereus* form distinct clusters while the other taxa group to form one cluster. For the microsatellite dataset, when K=5 *S. portenkoi* and *S. jacksoni* form their own distinct clusters, while *S. ugyunak* and *S. camtschatica* are still assigned as a single group, although two *S. camtschatica* individuals group very closely with *S. jacksoni*. For the SNP dataset, when K=5 *S. portenkoi* and *S. ugyunak* form more distinct clusters, although both still share some genetic similarity.

Phylogenetic analysis: mtDNA Cytb gene

The two methods implemented to infer phylogenies based on the Cytb dataset showed similar general relationships, but the Bayesian multispecies coalescent method allowed for higher resolution and stronger cladal support than the maximum likelihood method (Figs. 2.8, 2.9). The best nucleotide substitution model was TN93 with rate heterogeneity and invariant sites. Both Bayesian and ML methods confirm *S. cinereus* as an outgroup in relation to the remaining Beringian shrews, and both provide strong support for the monophyly of *S. pribilofensis*. The ML tree supports monophyly of *S. portenkoi*, but renders *S. ugyunak*, *S. camtschatica*, and *S. jacksoni* as polyphyletic (Fig. 2.8). These relationships are consistent and clearer based on the Bayesian tree (Fig. 2.9). Among Beringian taxa, *S. pribilofensis* is the first

taxon to diverge and forms a well-supported and reciprocally monophyletic group with respect to other Beringian shrews, although it exhibits little intra-specific divergence. The monophyly of *S. portenkoi* is also supported, but again, *S. jacksoni*, *S. camtschatica*, and *S. ugyunak* remain polyphyletic. Following initial differentiation of *S. pribilofensis*, divergence of the remaining Beringian shrews is shallow and remains unresolved.

Phylogenetic analysis: SNPs

Phylogenetic analysis of the SNP dataset provided increased support of taxon relationships observed based on Cytb data. The ML tree produced in RAxML and the coalescent species tree inferred with SVDquartets support *S. cinereus* as the sister taxon in relation to the other Beringian shrews (Figs. 2.10, 2.11). Again, *S. pribilofensis* diverged from the other two members of the Beringian clade (*S. ugyunak* and *S. portenkoi*) earliest, and is strongly supported as a monophyletic clade. Both of these methods resolve *S. ugyunak* and *S. portenkoi* as reciprocally monophyletic, although there is higher support for the monophyly of *S. ugyunak* in the SVDquartets coalescent species tree than from the ML tree. The ML tree indicates extremely low intra-specific differentiation within *S. pribilofensis* after initial divergence from *S. ugyunak* and *S. portenkoi*. The Bayesian coalescent species tree inferred with SNAPP supports these taxon relationships with maximum posterior probability (Fig. 2.12).

Species delimitation

Bayes Factor species Delimitation strongly supported the four-species model, in which *S. cinereus*, *S. pribilofensis*, *S. portenkoi*, and *S. ugyunak* are each recognized as a distinct species (Table 2.7). Using the marginal likelihood estimation (MLE) of each model, I compared the four-species model with each of the two alternate models (three-species and two-species; Table

2.3) using Bayes factors ($BF = 2 \times (MLE_1 - MLE_2)$). A positive BF value indicates support for model 1 while a negative BF value indicates support for model 2. The strength of support of Bayes factor comparison of competing models can be interpreted as: $0 < BF < 2$ is undecisive, $2 < BF < 6$ is mild support, $6 < BF < 10$ is strong support, and $BF > 10$ is decisive (Kass & Raftery, 1995; Leaché & Bouckaert, 2018). The strength of support for the four-species model in comparison to both of the alternate models is very decisive, with Bayes factors much greater than 10 for both comparisons.

Discussion

My analysis of ~11,000 SNPs, in combination with 17 microsatellites and the mtDNA cytochrome b gene, has provided higher resolution of species relationships within the Beringian clade of the *Sorex cinereus* species complex. The inclusion of three genetic datasets that are complimentary through their unique strengths constitutes a rigorous modern genomics approach to phylogenetic investigations. Microsatellites evolve very quickly, allowing for analysis of population-level differentiation over short time-scales. Cytb sequences provide a matrilineal history of diversification and can be resolvable over recent to deeper (>100kyr) timeframes. By genotyping thousands of SNP loci, it is possible to assess fine-scale evolutionary divergence and more accurately estimate phylogenies. Given that my study species are all closely related, and my primary question is if we can designate distinct species after comparatively brief divergence, incorporating loci with different levels of resolution is critical. As such, this study represents the most comprehensive assessment to-date of previously ambiguous species relationships within the Beringian clade of the *Sorex cinereus* species complex.

Despite the current taxonomic recognition of five shrews of the Beringian clade included in my thesis (*S. camtschatica*, *S. portenkoi*, *S. pribilofensis*, *S. jacksoni*, *S. ugyunak*), species-level status was questionable based on the existing literature (Hope et al., 2012). The observed divergence between these taxa was estimated to have occurred within the last 50 ky, through allopatry as a result of climate cycling through the late Quaternary. My findings provide insight that reflects strong evolutionary responses of these species to past environmental changes. First, there is clear differentiation between *S. pribilofensis* and its sister taxa across all locus-sets. The level of divergence is surprising given that differentiation has proceeded only since the most recent inundation of the Bering Isthmus, an incredibly short evolutionary timescale. In accordance with the likely order of geographic isolation, *S. pribilofensis* was the first to diverge from the rest of the high-latitude Beringian shrews, but also appears to have retained little to no intra-specific diversity. Based on Cytb sequences, relationships remain polyphyletic within the remaining shrews of the Beringian clade. However, phylogenetic analysis of the SNP dataset does resolve reciprocal monophyly of *S. ugyunak* and *S. portenkoi*, pointing again to allopatric divergence since inundation of the Bering Isthmus. While divergence has been recent, and ambiguous even through rapidly evolving microsatellite loci, my phylogenomic analysis and species delimitation methods using SNP data indicate multiple distinct evolutionary trajectories and rapid speciation within this species complex, lending support for current systematic relationships and providing an example of rapid mammalian evolution in response to late Quaternary climate change.

High genetic differentiation

Multiple genetic differentiation statistics (F_{ST} , G_{ST} , Jost's D, percent haplotype divergence), implemented for three types of genetic marker (microsatellites, mitochondrial gene,

nuclear SNPs), generally showed the same trends in species differentiation. Consistent with phylogenetic position as an outgroup to the members of the Beringian clade, *S. cinereus* showed highest differentiation from the other shrews based on the Cytb and SNP datasets. *Sorex pribilofensis* was the next most differentiated given these datasets, again consistent with phylogenetic results showing it to be the first to divergence from the rest of the Beringian clade. Of interest, genetic differentiation statistics for the microsatellite dataset were higher for *S. pribilofensis* than *S. cinereus*. This likely reflects the rapid rate of evolution of microsatellites from two perspectives. For *S. pribilofensis*, small populations in comparison to the mainland species, as a result of small island size, has allowed for rapid evolution coupled with loss of most diversity across microsatellites; this has led to very few alleles in common with other taxa. Simultaneously, *S. cinereus* has been isolated from other Beringian shrews for a full glacial cycle (~100 kyr), and given very high population size, likely has retained microsatellite diversity while also experiencing homoplasy due to reverse mutational steps (indels) of tandem repeats. If true, this phenomenon may give the impression of less differentiation even though shared microsatellite alleles are not derived from a common ancestor, and is one reason why microsatellites are generally not reliable for distinguishing between species. Regardless of the relative degree of genetic divergence, the values of genetic differentiation statistics for both *S. pribilofensis* and *S. cinereus* are generally high and on par with other estimates of species-level differentiation within mammals (Funk et al., 2016; Roy et al., 1994).

The other members of the Beringian clade (*S. camtschatica*, *S. portenkoi*, *S. jacksoni*, *S. ugyunak*) showed much less genetic differentiation, on par with estimates of mammalian differentiation between sub-species (Funk et al., 2016), and even isolated populations of the same species (Giglio et al., 2020; White & Searle, 2007). However, these estimates from the

literature came between sub-species and populations with extreme geographic isolation on multiple islands, or across heavily fragmented (patchy) landscapes. While these Beringian shrews have diverged in allopatry, it is likely that large ranges on the continents (and the relatively large Saint Lawrence Island for *S. jacksoni*) in combination with more recent geographic isolation have contributed to a slower accumulation of genetic differences in comparison to *S. pribilofensis* and *S. cinereus*. When taken in tandem with knowledge of recent allopatry, these levels of genetic differentiation, while not as high as noted for other species, point to independently evolving species on separate evolutionary trajectories.

The observed genetic differentiation statistics are corroborated and clearly visualized by the results of DAPC and Structure analyses. The most highly differentiated species for both clustering analyses and from both microsatellite and SNP datasets are *S. cinereus* and *S. pribilofensis*. In fact, for DAPC analysis of both genetic markers, the first two principle components relate almost entirely to the distinction of *S. cinereus* and *S. pribilofensis*, leaving little apparent differentiation between the remaining shrews. However, the third (and fourth, for the microsatellites) principle components show differentiation between *S. portenkoi*, *S. ugyunak*, *S. jacksoni*, and *S. camtschatica*, although divergence is minimal compared with *S. cinereus* and *S. pribilofensis*.

This is mirrored by Structure results, which find K=3 to be the most likely number of genetic clusters for both the microsatellite and SNP datasets, although K=5 also describes relevant differentiation. This again likely reflects relatively little genetic differentiation between *S. portenkoi*, *S. ugyunak*, *S. jacksoni*, and *S. camtschatica* in comparison to *S. cinereus* and *S. pribilofensis*. While K=3 best describes the levels of genetic differentiation within this complex, there still appears to be relevant structure beyond only three genetic clusters. Based on raw

likelihood scores, K=5 has the highest likelihood for both the microsatellite and SNP datasets, and allows identification of *S. portenkoi* and *S. jacksoni* as distinct clusters. Interestingly, *S. ugyunak* and *S. camtschatica* remain grouped together based on the microsatellite dataset, although this similarity may reflect a low sample size for *S. camtschatica*; additional sampling and/or genomic sequencing of *S. camtschatica* could better capture relevant genetic differentiation between these species. Overall, my findings indicate unambiguous genetic differentiation of *S. pribilofensis* from other taxa in comparison to less, but still relevant, genetic differentiation between the remaining shrews of the Beringian clade.

Rapid speciation

Phylogenetic analysis of the SNP dataset corroborates and provides increased resolution for the species relationships observed using the Cytb gene in this thesis and in previous studies (Demboski & Cook, 2003; Hope et al., 2012). When comparing the two methods used in my thesis to estimate phylogenies based on the Cytb gene, I consider the Bayesian phylogeny inferred with BEAST2 to be more accurate, given its use of coalescent theory coupled with biologically meaningful priors (e.g. mutation rate). The Cytb coalescent gene tree largely aligns with the phylogenies inferred from SNP data, providing insight for the matrilineal relationships of these shrew species that cannot be inferred from the SNP dataset. Based on all three phylogenies inferred with the SNP dataset, there is strong evidence for the initial split of *S. pribilofensis* from the other shrews of the Beringian clade. Given estimated divergence times by Hope et al., (2012), this split is congruent with the most recent formation of Saint Paul Island, ~14 kya. There appears to be no intra-specific differentiation across the island, as sub-specific clade structure within *S. pribilofensis* is virtually non-existent in any phylogeny. This points to a

history of rapid allopatric divergence coupled with severe loss of diversity across a very short evolutionary timescale.

Time calibrated phylogenetic inference shows diversification between the remaining members of the Beringian clade (*S. portenkoi*, *S. ugyunak*, *S. jacksoni*, *S. camtschatica*) occurred after the initial split of *S. pribilofensis*, lining up with the expectation that allopatric divergence in response to inundation of the Bering Isthmus led to differentiation. The formation of Saint Lawrence Island and the formation of the Bering Strait occurred more recently (~11 kya) than the formation of Saint Paul Island; in combination with larger available ranges (even on Saint Lawrence, which is ~65 times larger than Saint Paul), a lesser degree of divergence is expected. This is reflected by the coalescent *Cytb* gene tree, which suggests *S. ugyunak*, *S. camtschatica*, *S. jacksoni*, and to a certain extent *S. portenkoi* remain polyphyletic. Because mitochondrial DNA is haploid and maternally inherited, this likely reflects incomplete lineage sorting, as a few ancestral haplotypes remain shared between these species. However, the monophyly of both *S. portenkoi* and *S. ugyunak* is recovered by the SNP phylogenies. As the SNP dataset reflects a fairly random sample of the entire nuclear genome, the monophyly of these taxa suggests substantial realized divergence despite recent allopatry and shared *Cytb* haplotypes. Furthermore, I explicitly tested alternative hypotheses of species limits through Bayes Factor species Delimitation, the results of which strongly supported the species designations of *S. pribilofensis*, *S. ugyunak*, and *S. portenkoi* over alternate species delimitation models. As the coalescent *Cytb* gene tree largely reflects the SNP phylogenies, just at lower resolution, it is likely that the shared haplotypes between *S. jacksoni* and *S. camtschatica* also reflect incomplete lineage sorting and that the monophyly of these taxa would be recovered based on nuclear SNPs. However, without this data their monophyly cannot be recovered conclusively.

Distinct evolutionary trajectories

Analysis of multiple independent genetic datasets, under phylogenetic and population genetic frameworks, points towards distinct evolutionary trajectories for each currently recognized species of the Beringian clade included in this thesis, despite varying levels of genetic differentiation and phylogenetic divergence. Under the phylogenetic species concept, *S. pribilofensis*, *S. ugyunak*, and *S. portenkoi* represent distinct species, as the SNP data shows they are reciprocally monophyletic and independently evolving. Given the near monophyly of *S. camtschatica* and *S. jacksoni* based on the coalescent Cytb gene tree, and similar, if not higher, levels of differentiation in comparison to *S. ugyunak* and *S. portenkoi*, I suggest that these taxa continue to be recognized as distinct species as well. There remains the argument that were any of these five species to come in contact with each other they would be able to interbreed, as is often the case when species evolve in allopatry (Coyne & Orr, 1989; Matute & Cooper, 2021). If this were to be the case, then under the biological species concept it could be argued that these taxa don't represent distinct species, but only subspecies or independent populations of a single species. However, given the degree of geographic isolation (across continents and highly isolated islands), these terrestrial species are not naturally able to come in contact with each other unless sea levels are again lowered during a glacial phase (increasingly unlikely given anthropogenic global warming and associated rising sea levels; Stocker et al., 2013), or individuals are introduced across barriers either intentionally or otherwise. The exception to this is *S. portenkoi* and *S. camtschatica*, which are both distributed in Far East Asia. In the case of these two species, however, their distinctness remains clear despite the potential for sympatry and interbreeding; *S. portenkoi* remains monophyletic across both Cytb and SNP phylogenies, and *S. camtschatica* appears to be more closely related to *S. ugyunak* (from Alaska) based on both population genetic and phylogenetic analyses. Because of such strict geographic isolation between the majority of

the species in the Beringian clade, any potential ability of these species to interbreed is largely irrelevant, as each species continues on its unique and independent evolutionary trajectory.

Given all the evidence, I recognize Beringian shrews of the *cinereus* complex as independent species. Reciprocal monophyly was recovered based on thousands of SNPs for the species analyzed, in accordance with the phylogenetic species concept, which describes species as monophyletic groups derived from a common ancestor and which possess a set of diagnosable synapomorphies (Baum & Donoghue, 1995). More importantly, even considering polyphyly of some taxa based on a mtDNA gene, these shrews are experiencing independent evolutionary trajectories. This satisfies the basic tenet of the unified species concept (deQuieros, 2005), which allows for situations, such as this, where the rigidity of more traditional species concepts falls short of accounting for complex biological realities.

Conclusions

There is increasing evidence that periodic climate cycling of the Quaternary has had a heavy influence on species distributions throughout the northern hemisphere (Avise et al., 1998; Hewitt, 2000; Hewitt, 2004). While early phylogeographic studies have shown divergence through refugia during glacial periods (Comes & Kadereit, 1998; Taberlet et al., 1998), this study provides an example of exceptionally rapid speciation in a mammalian system during the present interglacial (Holocene), and at maximum within the timeframe of the last glacial maximum (<20 kyr). In particular, I provide evidence for the rapid speciation and subsequent loss of most intra-specific variation within *S. pribilofensis* following recent island isolation. As a federally recognized endangered species, the findings of this study warrant its continued population monitoring and federal protection. In addition to allowing for timely identification of possible population declines, continued research on *S. pribilofensis* will provide further insight into not

only the dynamics of its own evolutionary trajectory, which will inform conservation practices, but on the evolutionary response of small-island endemics to past and current environmental change. A promising future line of research is whole genome sequencing, which is becoming increasingly tractable for non-model organisms (Ellegren, 2014; Jones & Good, 2016), and could shed further light on the evolutionary trajectory of *S. pribilofensis* in addition to providing insight into species' response to island isolation at functional regions of the genome. As biodiversity worldwide responds to accelerating anthropogenic climate change, knowledge of evolutionary responses to past dramatic environmental change will become increasingly imperative for effective conservation strategies (Colella et al., 2020; Hoffmann & Sgrò, 2011). This study provides evidence for rapid evolutionary change within a mammalian system in response to climate cycling of the late Quaternary and highlights the unique evolutionary trajectory of an endangered island endemic in the face of continued environmental change.

Figures

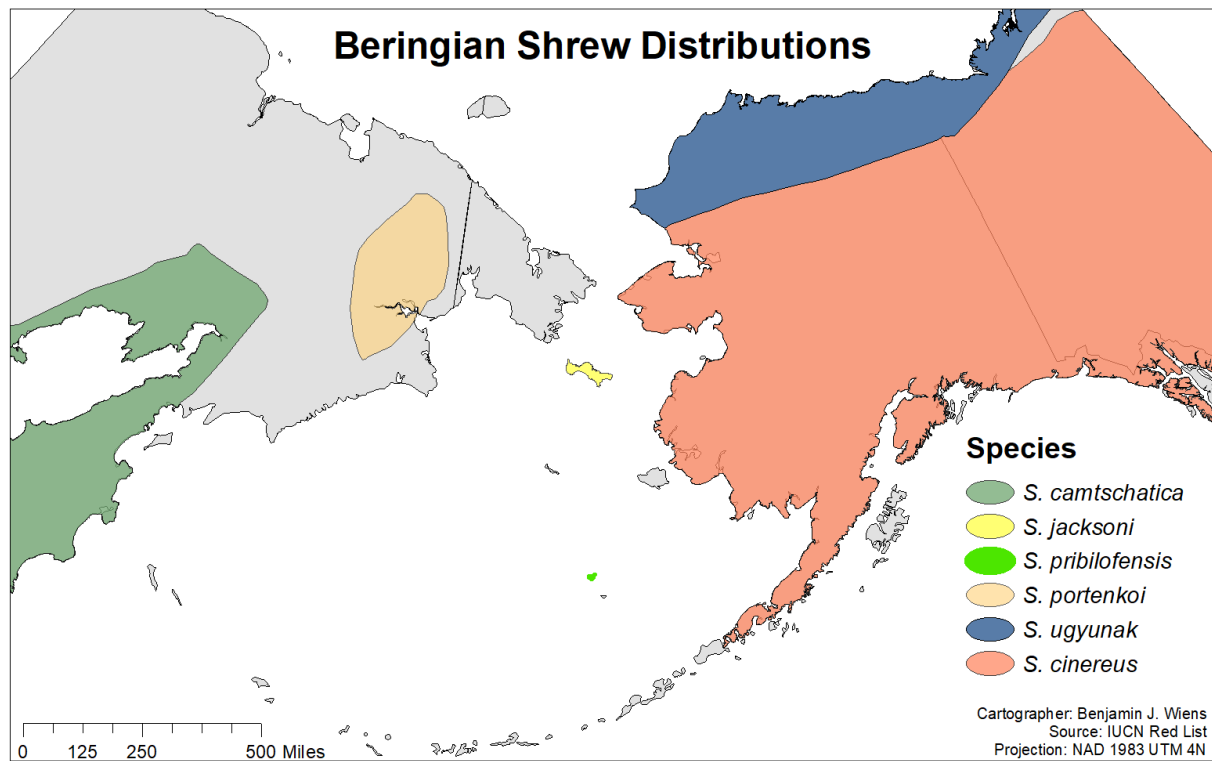


Figure 2.1 Distribution of shrews included in this study

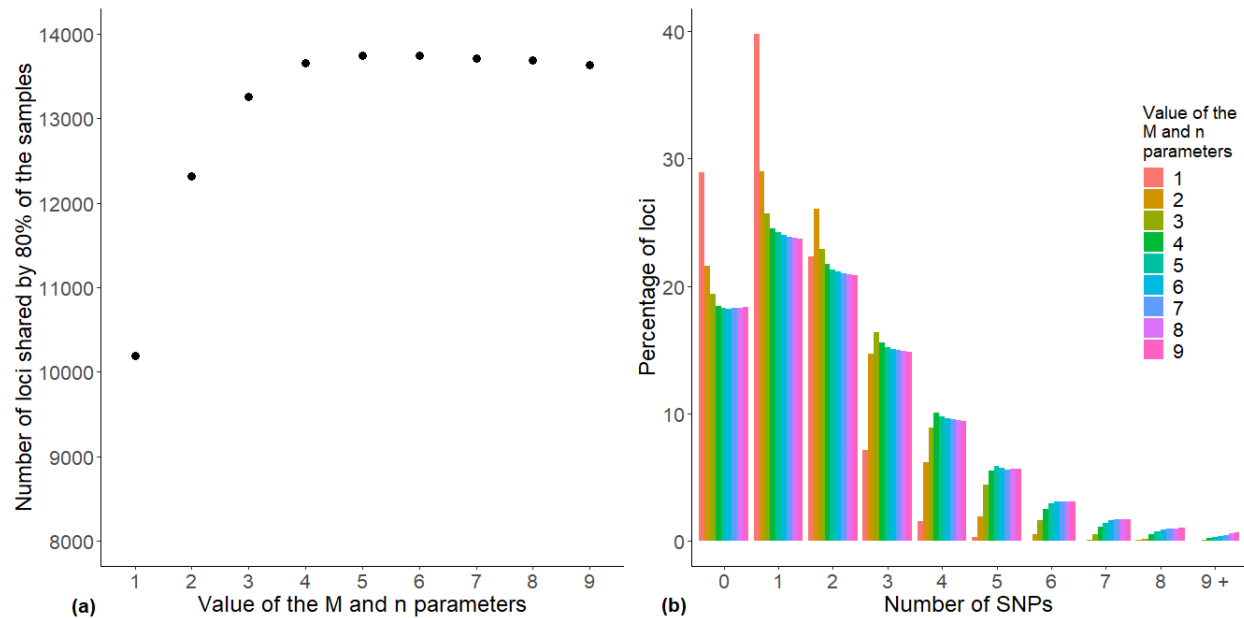


Figure 2.2 ddRADseq parameter selection

Selection of parameters for *denovo* assembly in Stacks v2.54. **(a)** Increasing the value of the M and n parameters, while holding the m parameter constant at 3, increases the number of identified loci until $M=n=4$. **(b)** The percentage of loci assembled, categorized by the number of SNPs per locus, plateaus around $M=n=4$. Before $M=n=4$, a high percentage of loci with 0-2 SNPs are assembled, but a low percentage of loci with 3-9+ are assembled. $M=n=4$ optimizes the number of loci assembled.



Figure 2.3 Read depth per SNP locus

Distribution of read depth per locus after initial quality filters for shrews included in SNP dataset. Average read depth=15.75. First three letters indicate species (SCI=*S. cinereus*; SPB=*S. pribilofensis*; SPO=*S. portenkoi*; SUG=*S. ugunak*), and following specimen numbers are coincident with Table A1.

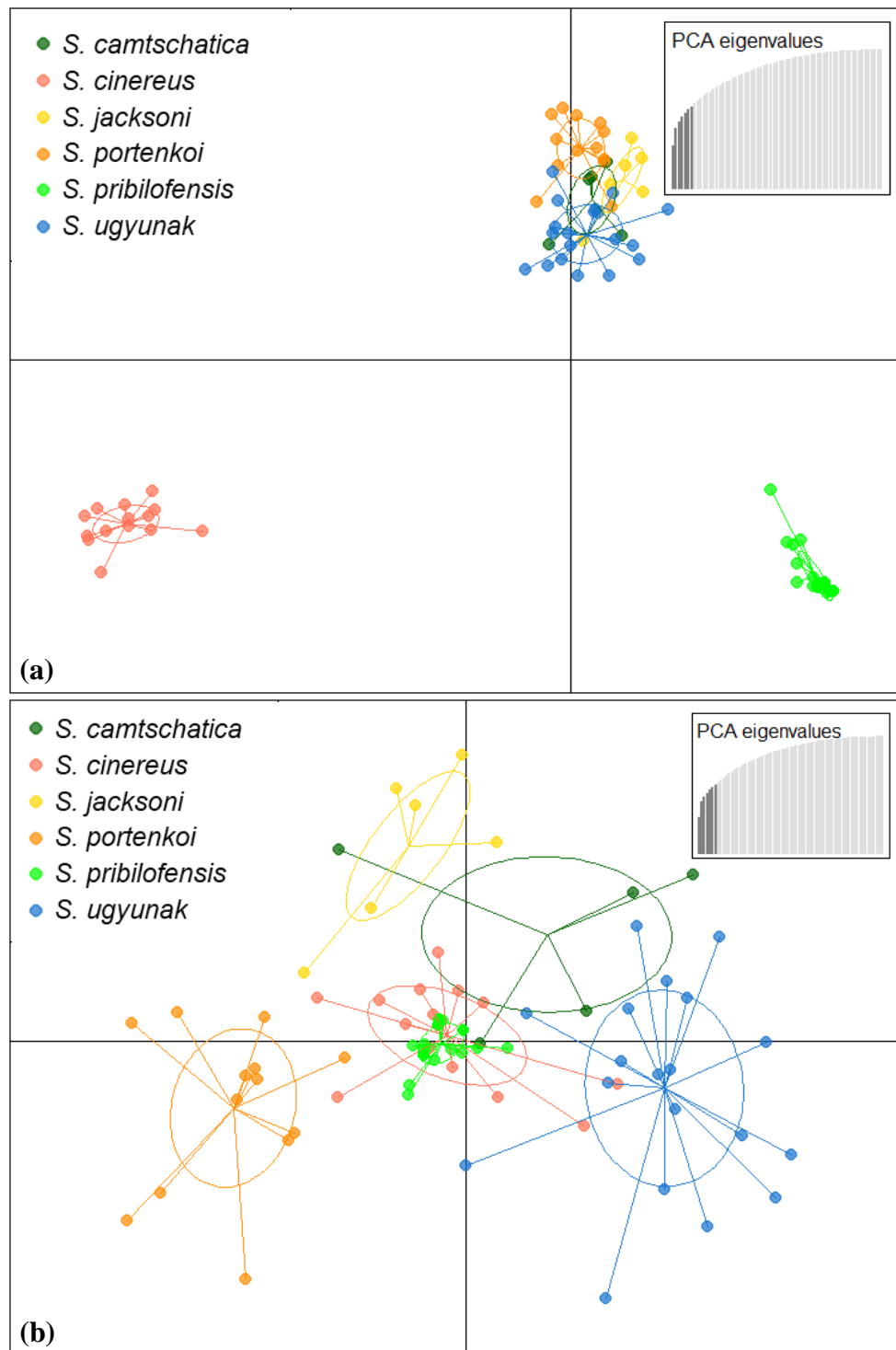


Figure 2.4 Microsatellite DAPC

DAPC scatter plots based on the microsatellite dataset. The cumulative variance explained by the 7 principle components retained for this analysis is shown in the top right of each plot. **(a)** DAPC using the first and second principle components; **(b)** DAPC using the third and fourth principle components.

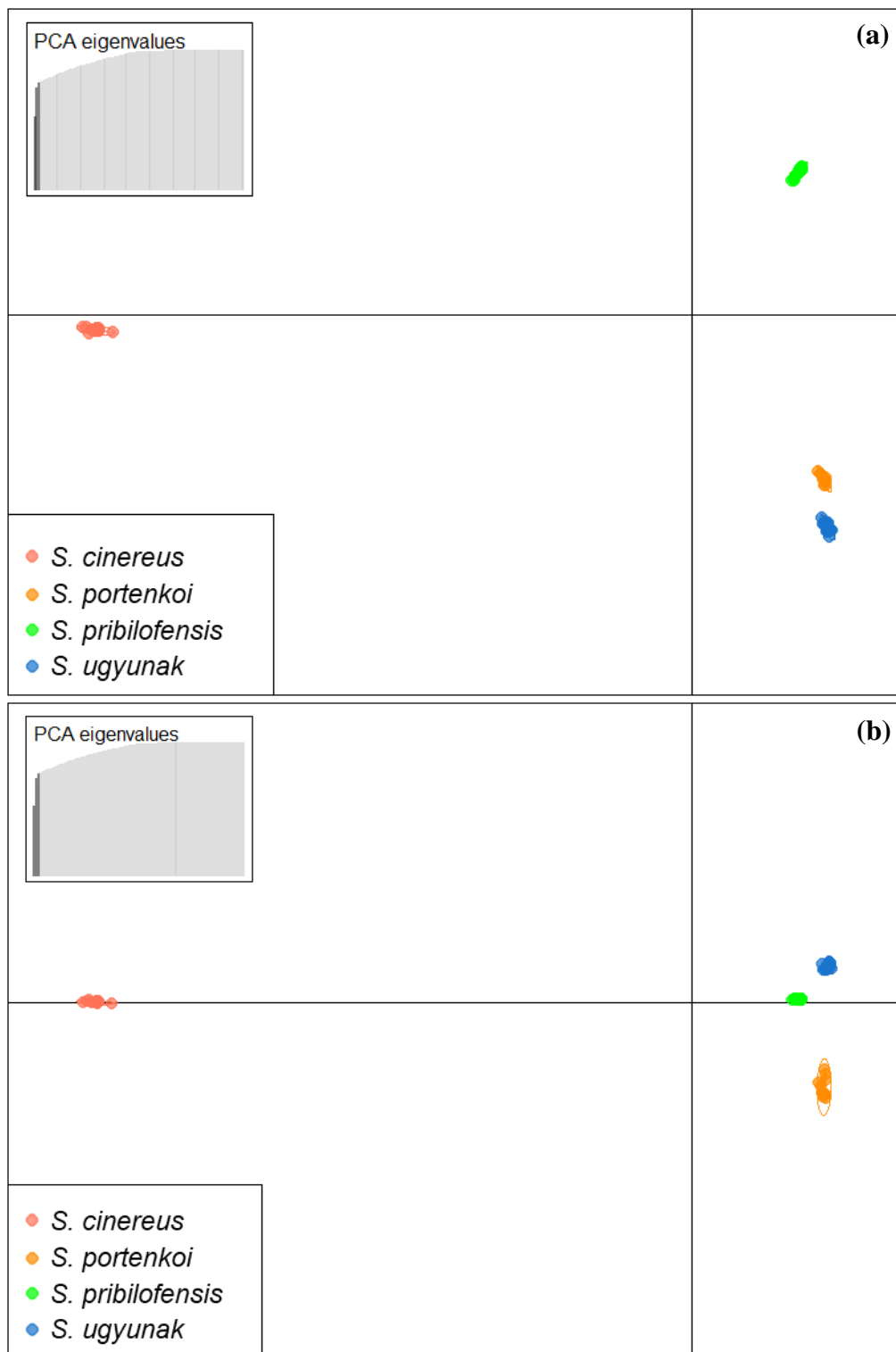


Figure 2.5 SNP DAPC

DAPC scatter plots based on the SNP dataset. The cumulative variance explained by the 3 principle components retained for this analysis is shown in the top left of each plot. **(a)** DAPC using the first and second principle components; **(b)** DAPC using the first and third principle components.

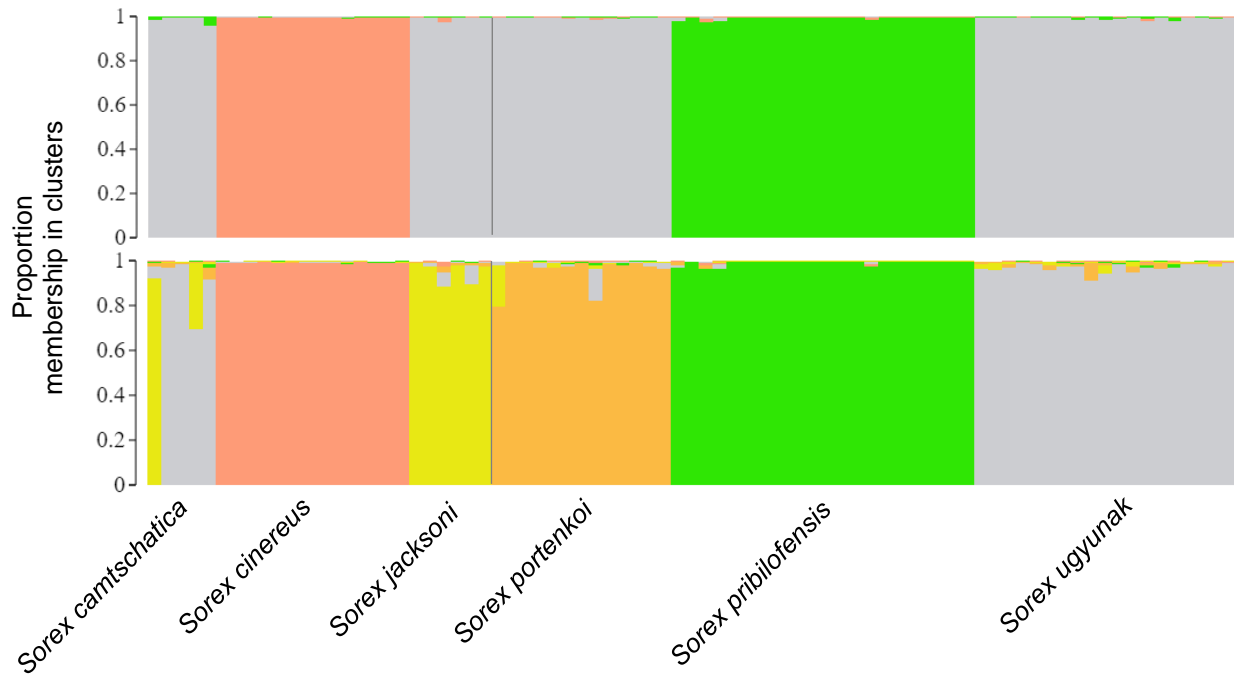


Figure 2.6 Microsatellite Structure plots

Structure plots based on the microsatellite dataset. Each color corresponds to a distinct genetic cluster, and each bar represents the proportion of each individual's genotype to each cluster. **(top)** Structure plot for most likely number of genetic clusters, $K=3$; **(bottom)** Structure plot for second most likely number of genetic clusters, $K=5$.

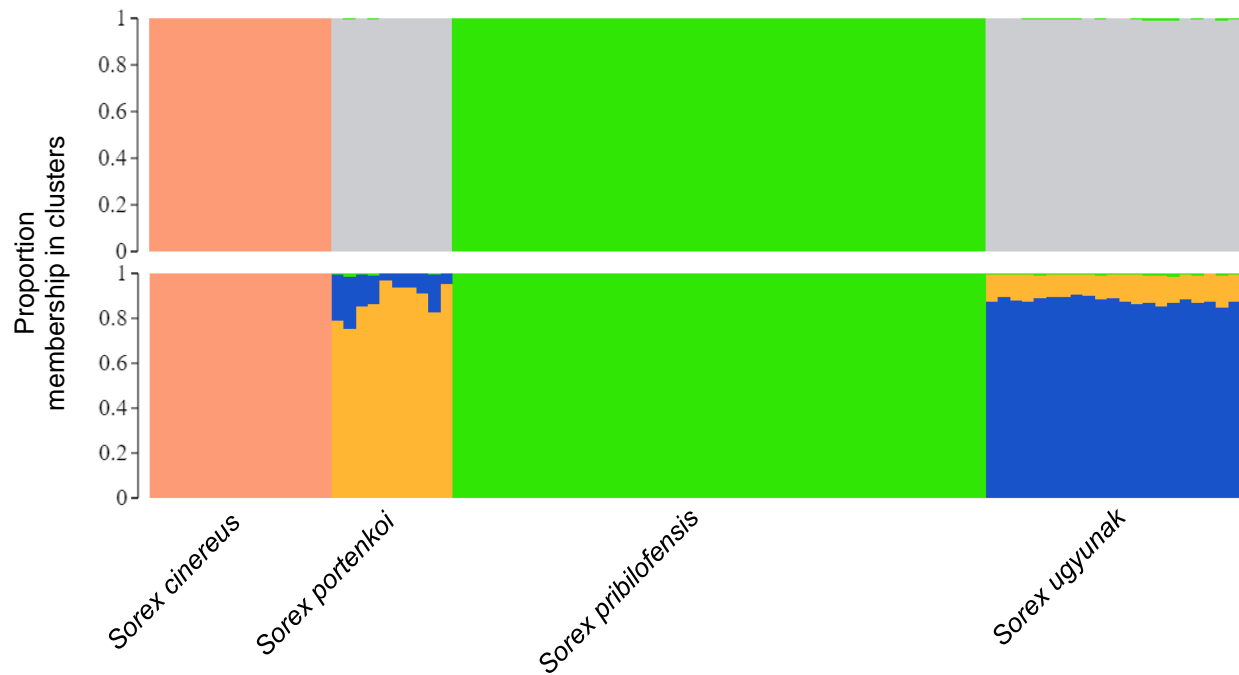


Figure 2.7 SNP Structure plots

Structure plots based on the SNP dataset. Each color corresponds to a distinct genetic cluster, and each bar represents the proportion of each individual's genotype to each cluster. **(top)** Structure plot for most likely number of genetic clusters, $K=3$; **(bottom)** Structure plot for second most likely number of genetic clusters, $K=5$. Note that for $K=5$, no measurable portion of any individual was assigned to the fifth cluster, so $K=5$ essentially reflected 4 clusters.

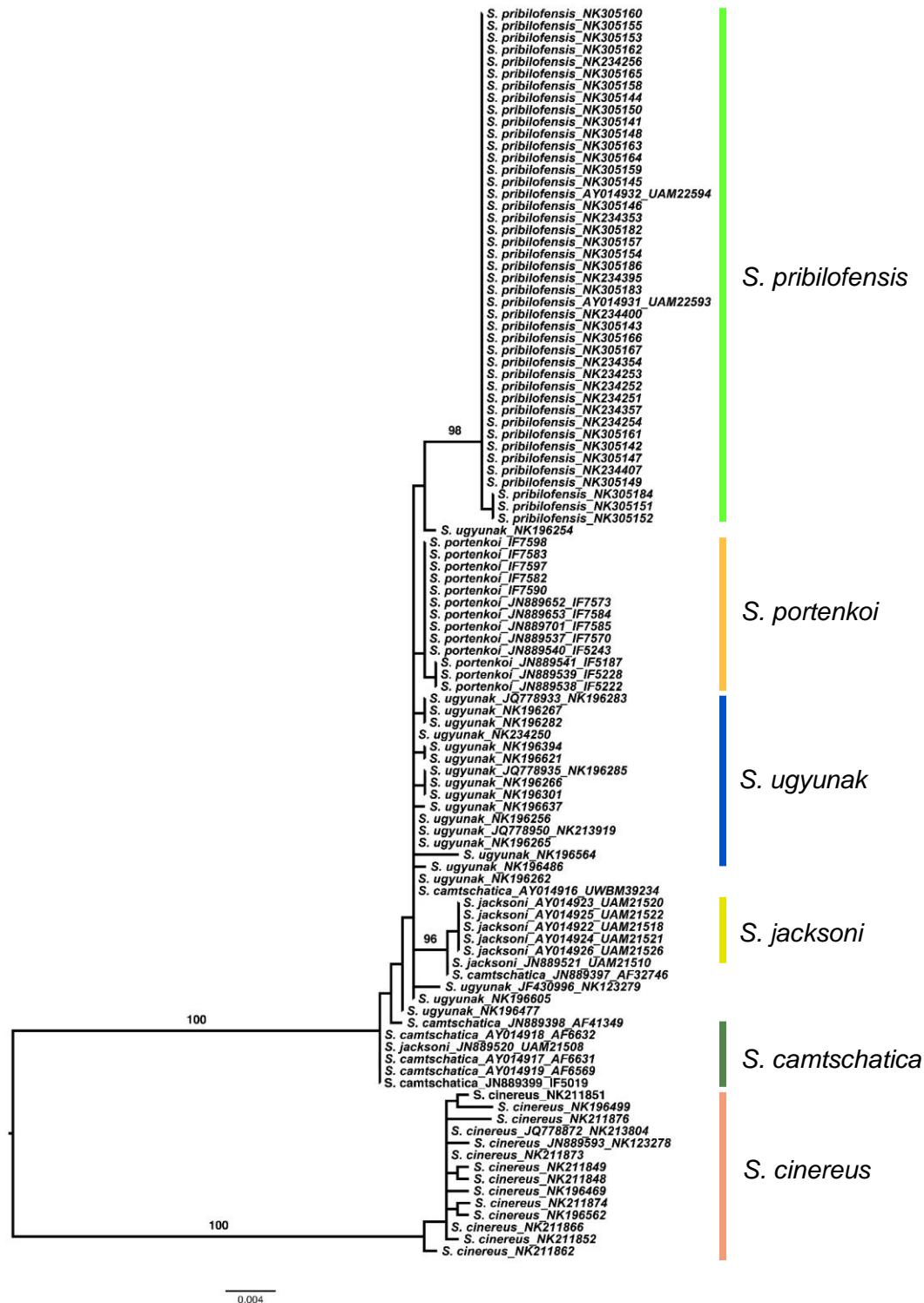


Figure 2.8 Cytb RAxML phylogeny

Maximum likelihood phylogeny produced with RAxML for the Cytb gene. Specimen numbers are coincident with Table A1. Branch labels indicate the percentage of bootstraps that support the subsequent clade. Bootstrap percentages <80 for minor clades are not shown and major unlabeled nodes are only weakly supported.

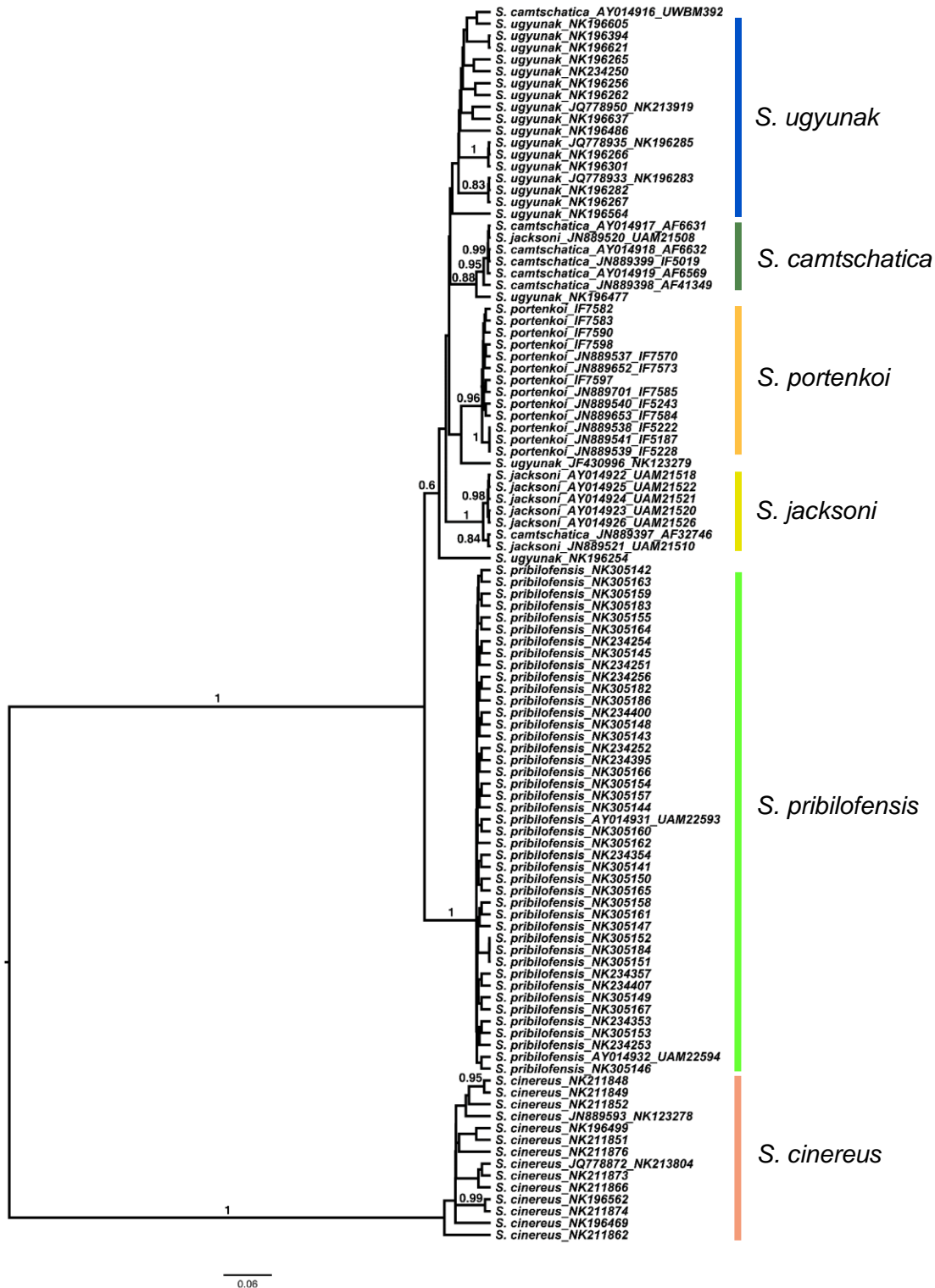


Figure 2.9 Cytb BEAST2 phylogeny

Bayesian coalescent phylogeny produced with BEAST 2 for the Cytb gene. Specimen numbers are coincident with Table A1. Branch labels indicate posterior probability support. Posterior probabilities <0.80 for minor clades are not shown and major unlabeled nodes are only weakly supported.

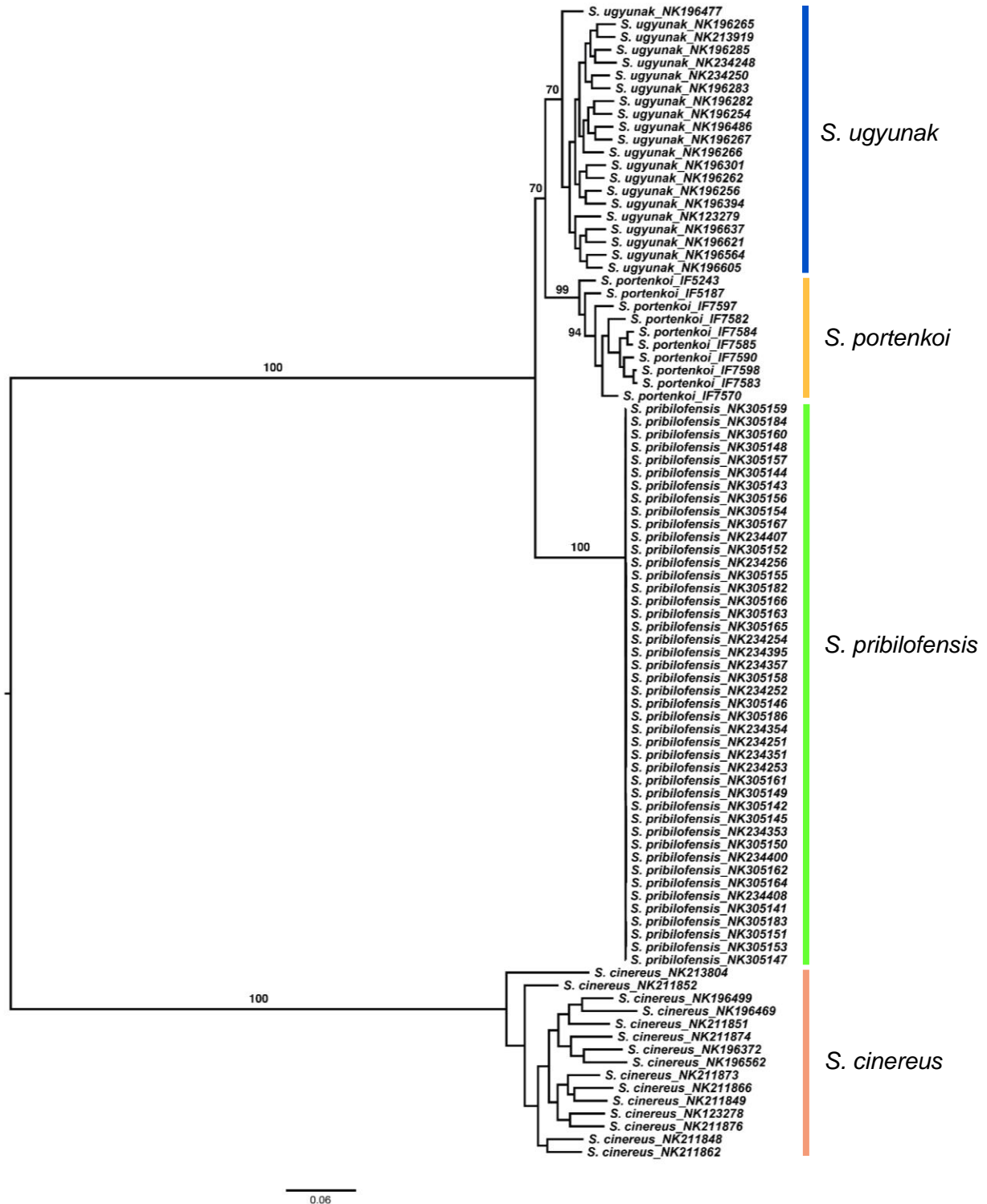


Figure 2.10 SNP RAxML phylogeny

Maximum likelihood phylogeny produced with RAxML for the SNP dataset. Specimen numbers are coincident with Table A1. Branch labels indicate the percentage of bootstraps that support the subsequent clade. Bootstrap percentages <80 for minor clades are not shown and major unlabeled nodes are only weakly supported.

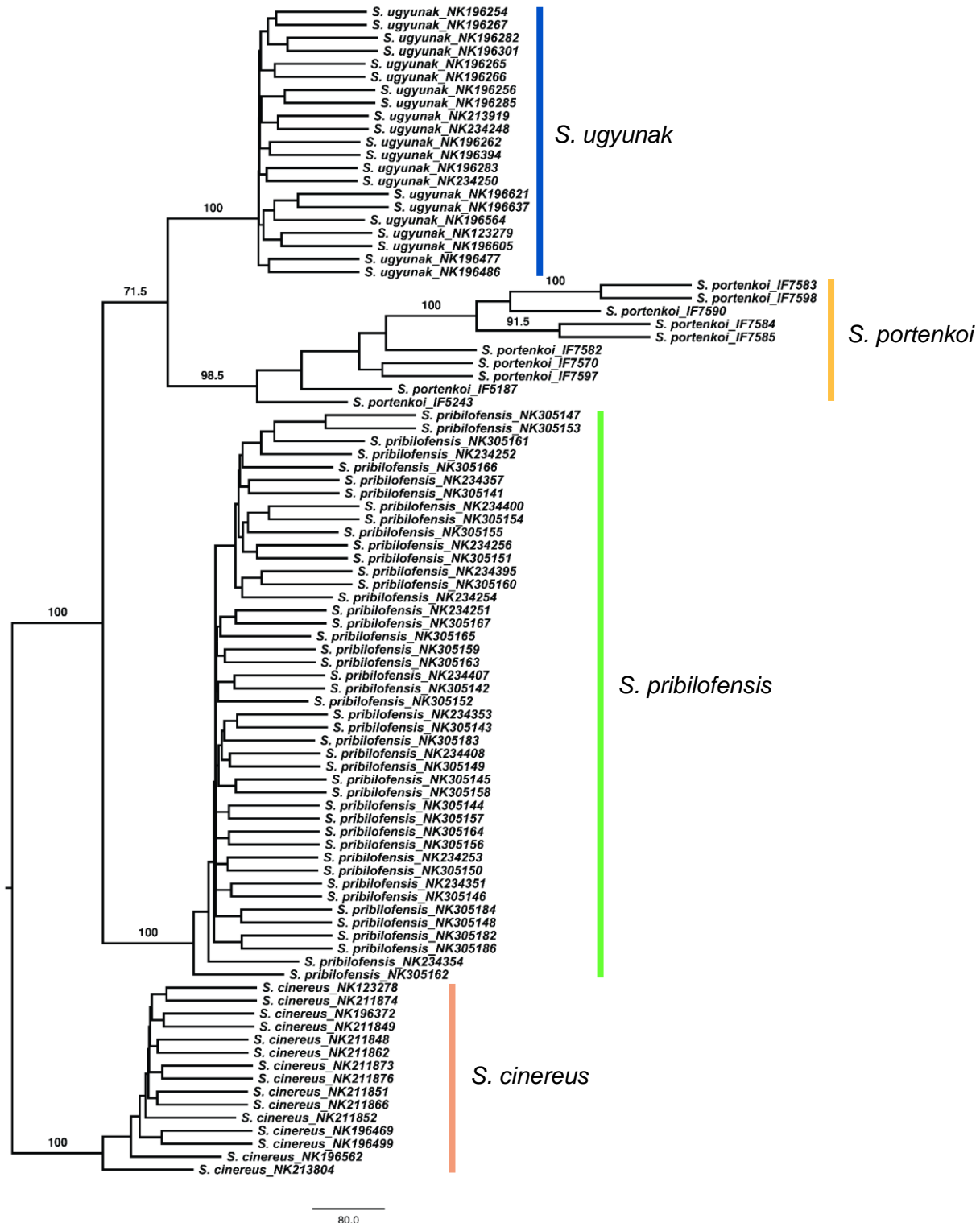


Figure 2.11 SNP SVDquartets phylogeny

Coalescent phylogeny produced with SVDquartets for the SNP dataset. Specimen numbers are coincident with Table A1. Branch labels indicate the percentage of bootstraps that support the subsequent clade. Bootstrap percentages <80 for minor clades are not shown and major unlabeled nodes are only weakly supported. Branch lengths correspond to bootstrap support.

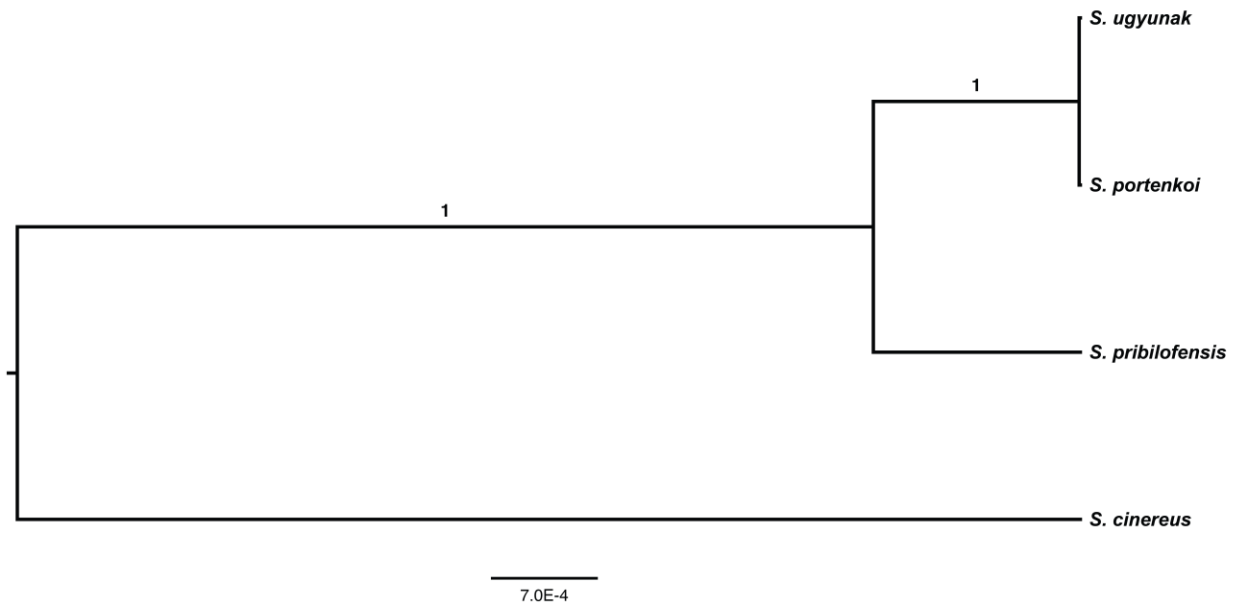


Figure 2.12 SNP Bayesian phylogeny

Bayesian coalescent species-tree phylogeny produced with SNAPP for the SNP dataset. Each node represents eight randomly selected individuals for that taxon. Branch labels indicate posterior probability support.

Tables

Table 2.1 Shrew sample sizes

Number of individuals of each species included in the three genetic datasets.

Species	Microsatellite	Cytochrome b	SNP
<i>Sorex camtschatica</i>	5	7	0
<i>Sorex cinereus</i>	14	14	15
<i>Sorex jacksoni</i>	6	7	0
<i>Sorex portenkoi</i>	13	13	10
<i>Sorex pribilofensis</i>	22	43	44
<i>Sorex ugyunak</i>	19	20	21
Total	79	104	90

Table 2.2 ddRADseq missing data

Average proportion of missing data per species for the SNP dataset after all quality filters.

Species	Average proportion of missing data per individual
<i>Sorex cinereus</i>	20.9987%
<i>Sorex pribilofensis</i>	5.3320%
<i>Sorex portenkoi</i>	5.4734%
<i>Sorex ugyunak</i>	5.4832%

Table 2.3 Species delimitation models

Species delimitation models tested through Bayes Factor species Delimitation.

Species model	Species 1	Species 2	Species 3	Species 4	Rational
Four-species	<i>S. cinereus</i>	<i>S. pribilofensis</i>	<i>S. portenkoi</i>	<i>S. ugyunak</i>	current taxonomy
Three-species	<i>S. cinereus</i>	<i>S. pribilofensis</i>	<i>S. portenkoi</i> + <i>S. ugyunak</i>		structure clustering results
Two-species	<i>S. cinereus</i>	<i>S. pribilofensis</i> + <i>S. portenkoi</i> + <i>S. ugyunak</i>			shallow divergence within Beringian clade

Table 2.4 Microsatellite genetic differentiation

Genetic differentiation measures based on the microsatellite dataset. **(a)** Pairwise F_{ST} values; **(b)** pairwise G_{ST} values; **(c)** pairwise Jost's D values. Species codes: SOCT= *S. camtschatica* SOCI= *S. cinereus* SOJA= *S. jacksoni* SOPO= *S. portenkoi* SOPB= *S. pribilofensis* SOUG= *S. ugyunak*

(a)	SOCT	SOCI	SOJA	SOPO	SOPB	SOUG
SOCT	-					
SOCI	0.256	-				
SOJA	0.110	0.352	-			
SOPO	0.091	0.259	0.151	-		
SOPB	0.538	0.593	0.659	0.557	-	
SOUG	0.028	0.239	0.154	0.099	0.487	-

(b)	SOCT	SOCI	SOJA	SOPO	SOPB	SOUG
SOCT	-					
SOCI	0.153	-				
SOJA	0.061	0.218	-			
SOPO	0.051	0.150	0.083	-		
SOPB	0.373	0.423	0.483	0.386	-	
SOUG	0.019	0.137	0.087	0.053	0.323	-

(c)	SOCT	SOCI	SOJA	SOPO	SOPB	SOUG
SOCT	-					
SOCI	0.814	-				
SOJA	0.188	0.872	-			
SOPO	0.232	0.808	0.270	-		
SOPB	0.673	0.928	0.695	0.759	-	
SOUG	0.092	0.809	0.314	0.269	0.639	-

Table 2.5 Cytb genetic differentiation

Genetic differentiation measures based on the Cytb dataset. **(a)** Pairwise F_{ST} values; **(b)** percent haplotype divergence values. See Table 2.4 for species codes.

(a)	SOCT	SOCI	SOJA	SOPO	SOPB	SOUG
SOCT	-					
SOCI	0.802	-				
SOJA	0.465	0.852	-			
SOPO	0.507	0.889	0.551	-		
SOPB	0.668	0.889	0.715	0.801	-	
SOUG	0.401	0.847	0.438	0.061	0.538	-

(b)	SOCT	SOCI	SOJA	SOPO	SOPB	SOUG
SOCT	-					
SOCI	5.57	-				
SOJA	0.44	5.74	-			
SOPO	0.37	5.82	0.41	-		
SOPB	0.89	5.82	1.01	0.69	-	
SOUG	0.34	5.67	0.31	0.23	0.67	-

Table 2.6 SNP genetic differentiation

Genetic differentiation measures based on the SNP dataset. **(a)** Pairwise F_{ST} values; **(b)** pairwise G_{ST} values; **(c)** pairwise Jost's D values. See Table 2.4 for species codes.

(a)	SOCI	SOPO	SOPB	SOUG
SOCI	-			
SOPO	0.737	-		
SOPB	0.817	0.684	-	
SOUG	0.721	0.223	0.59	-

(b)	SOCI	SOPO	SOPB	SOUG
SOCI	-			
SOPO	0.586	-		
SOPB	0.692	0.52	-	
SOUG	0.565	0.127	0.419	-

(c)	SOCI	SOPO	SOPB	SOUG
SOCI	-			
SOPO	0.402	-		
SOPB	0.428	0.107	-	
SOUG	0.395	0.034	0.090	-

Table 2.7 Species delimitation

Comparison of alternate species models using Bayes Factor species Delimitation.

MLE=marginal likelihood estimate, BF=Bayes factor, calculated as $BF = 2 \times (MLE_1 - MLE_2)$.

Positive BF values greater than 10 indicate strong support for Model 1.

Model 1	Model 2	MLE₁	MLE₂	BF
Four-species	Three-species	-55,331.93	-59,614.98	8,566.10
Four-species	Two-species	-55,331.93	-70,686.30	30,708.74

Chapter 3 - Genetic drift drives differentiation of an endangered insular shrew (*Sorex pribilofensis*)

Introduction

Islands are hotspots of biodiversity and endemism. Marine islands make up only 5.3% of the world's total land area, but harbor a disproportionately large amount of the world's species (Tershy et al., 2015). However, island species are also vulnerable to anthropogenic environmental changes, which present multiple stressors, including invasive species, habitat loss, climate change, industrial and urban development, disease, and overexploitation (Leclerc et al., 2018; Russell & Kueffer, 2019). This has led to the disproportionate extinction of island species, accounting for 61% of all documented extinctions (Tershy et al., 2015). For those island species that persist today, half are recognized as at some risk of extinction, and island species account for 37% of all critically endangered species (Leclerc et al., 2018; Tershy et al., 2015). The biodiversity of land-bridge islands are particularly vulnerable to current global change, as a result of their unique evolutionary ecology stemming from fragmentation followed by long-term isolation (Stuart et al., 2012). Because land-bridge island populations were once part of a more broadly distributed community across some mainland areas prior to island isolation, their natural history is distinct from that of populations occurring on oceanic islands, which were never connected to a mainland. While the evolutionary history of oceanic island biodiversity is one of colonization and new community interactions, land-bridge island biodiversity instead has a history of ongoing faunal and floral relaxation (species loss) as communities adjust to new biogeographic equilibria following initial isolation (Lomolino, 2000).

The same evolutionary processes that lead to high levels of endemism and unique biodiversity connections on land-bridge islands also leave these species vulnerable. From an

evolutionary standpoint, these dynamics include divergent selection, genetic drift, and potential gene flow among isolated land-masses. For non-volant island species, both divergent selection and genetic drift may often act unimpeded by gene flow with mainland populations, due to a lack of immigration. In the absence of gene flow, both divergent selection and genetic drift can be expected to contribute to rapid divergence and eventual speciation of land-bridge island populations following initial isolation (Fisher, 1958). Therefore, an understanding of how these very different evolutionary mechanisms have shaped insular faunal assemblages in the past will be critical for predicting their responses to ongoing environmental change.

Divergent selection facilitates genetic differentiation between populations through adaptation to local conditions (Fisher, 1958). This allows for adaptive evolution of land-bridge island populations in response to distinct island environments. Adaptive evolution of isolated taxa is evident through many striking examples (e.g. finches, Grant, 2002; anole lizards, Losos, 1998; cichlid fishes, Terai et al., 2006). Genetic divergence due to local adaption to island environments is generally expected, as marine islands in general are environmentally distinct from the mainland (Weigelt et al., 2013). A common driver of this distinctness is climate, as island climates can be strongly influenced by their position to ocean currents, as well as by elevation and topography (Spalding et al., 2007). While adaptation to island climates promotes divergence and leads to the unique biodiversity of islands, it also leaves island species at risk if the climate were to rapidly change from specific historic conditions, because island species cannot respond by range-shifting, and must instead quickly adapt or persist through plasticity if they are to survive. Island environments are also distinctive in that the biotic community is typically less diverse than on mainland areas through the process of faunal (or floral) relaxation (Brown, 1978; Lomolino et al., 1989; Patterson, 1987). Species adapt to a lifestyle of interactions

with often only a small subset of more diverse mainland communities, and these interactions are often generalized and mutualistic (Kaiser-Bunbury et al., 2010). However, evolution in the presence of only few predators, competitors, and pathogens leaves island endemics vulnerable to invasive species, which can quickly lead to extinction of native species upon introduction to island systems. The process of divergent selection thereby facilitates rapid speciation on land-bridge islands, but can leave species vulnerable to rapid environmental changes.

Genetic drift can be just as powerful as divergent selection in the evolution of land-bridge island taxa (Funk et al., 2016; Jordan & Snell, 2008; Prentice et al., 2017; Toczydlowski & Waller, 2019). This is because the strength of genetic drift is inversely related to effective population size (N_e), which for a number of reasons is often small for land-bridge island taxa (Charlesworth, 2009; Frankham, 1997a). First, founder effects of the initial isolated population, as a subset of the mainland population, can lead to a small N_e through low genetic variation in the initial gene pool. Additionally, the relatively small size of islands compared to the mainland often results in declines in census size following isolation, which can reduce genetic variability through repeated and potentially large fluctuations in census size and subsequent genetic bottlenecks (Nei et al., 1975). So, due to small effective population sizes through founder effects, bottlenecks, and/or large fluctuation in census size, genetic drift can act strongly in island populations (Woolfit & Bromham, 2005). This can result in the rapid fixation of existing alleles and an overall reduction in genetic variability, and possibly lead to non-adaptive divergence and even speciation from mainland populations (Frankham, 1997a; Funk et al., 2016). However, this can put island endemics at risk in the face of rapid environmental change, as genetic variability is the raw material through which adaptive change can occur. Genetic drift can thereby contribute to rapid divergence of land-bridge island populations, while also leaving them with reduced

genetic variability and limited ability to adapt to new environmental conditions (Agashe et al., 2011; Reed & Frankham, 2003).

The influences of genetic drift and divergent selection on the differentiation of island taxa are not independent of each other (Funk et al., 2016; Ohta, 1992). Given strong enough selective pressure, adaptive change can occur even when effective population sizes are relatively low (Clegg et al., 2002). Alternatively, when effective population sizes are exceedingly low, it is difficult for adaptive change to take place in the face strong genetic drift (Charlesworth, 2009). This means that divergent selection and genetic drift can simultaneously shape the differentiation of island taxa, and we can make predictions about which evolutionary force will predominate given island conditions. For taxa isolated on small islands, which experience no gene flow, have gone through bottlenecks, and/or experience large fluctuations in census size, we can predict that the contribution of genetic drift to observed differentiation is large relative to that of divergent selection. Due to low N_e in this scenario, selection only overcomes the effects of genetic drift for loci under extreme selective pressure (Ohta & Tachida, 1990). Alternatively, for taxa isolated on large islands, which may experience gene flow, have not gone through severe bottlenecks, and/or have steady census sizes, we can predict that divergent selection has been the predominant contributor to any observed differentiation (Heaney, 2000). Due to high N_e in this scenario, selection is effective even when selective pressure is small, and it is unlikely for alleles to drift far from their starting frequencies (Ohta & Tachida, 1990). So under this framework, knowledge of island conditions and species demographics can inform predictions about the predominant evolutionary force leading to the differentiation of land-bridge island taxa.

Islands in the Bering Sea represent a classic system of land-bridge island evolution. Through the Quaternary there has been a cycle of sea level oscillations which have alternately

connected Alaska and Siberia through the Bering Isthmus and then disconnected the continents leaving only island remnants of the land bridge (Bintanja et al., 2005). Glacial-interglacial cycling in the Beringian region has thereby greatly influenced species distributions through changes in sea level and the creation of isolated refugia. Throughout the Quaternary, glaciers and continent-wide ice sheets extended from polar regions as far south as the conterminous United States in North America, but Beringia remained largely ice-free during both glacial and interglacial phases. As such, Beringia acted as a single large refugium for terrestrial fauna and as a route for interchange between Alaska and Siberia during glacial periods (Brigham-Grette et al., 2003). During interglacial phases, multiple smaller insular refugia formed throughout the Bering Sea, such as at present. From an evolutionary perspective, these isolation-reconnection dynamics had dramatic impacts on the phylogeography and population genetics of terrestrial fauna (Sher, 1999).

Beringian climate cycling has heavily influenced the phylogeographic history of the *Sorex cinereus* complex of shrews (hereafter *cinereus* complex; Hope et al., 2012). Members of the *cinereus* complex are wide-ranging, occurring throughout most of North America and into Siberia (western Beringia). This species complex currently consists of 13 species of the subgenus *Otisorex* (Hutterer, R., 2005), broken up into Beringian (the focus of my thesis) and Southern clades (Table 3.1; van Zyll de Jong, 1991; Demboski & Cook, 2003). The phylogeography of the Beringian clade is best characterized by allopatric divergence, followed by range expansion and/or reduction, as a result of geographic isolation. Initial divergence within the Beringian clade occurred between the mid-latitude taxa and high-latitude taxa, when the ancestral form of the high-latitude taxa expanded northward into Beringia during the Sangamon interglacial (~130 kya), following glacial retreat. The ancestral form of the high-latitude taxa tracked the recession

of tundra habitat as boreal forests expanded north, which ultimately form a forest barrier between the high-latitude and mid-latitude taxa. The onset of the Wisconsin glacial (~74 kya) would have further isolated the ancestral form of the high-latitude taxa in the large Beringian refugium, and allowed expansion of shrews across the land bridge into Siberia. Since the last glacial maximum (LGM; ~20 kya), a warming climate and rising sea levels led to the isolation of populations of this ancestral form on St. Paul Island (~14 kya), St. Lawrence Island (~8 kya), Paramushir Island (~8-20 kya), and on either side of the Bering Strait (~11 kya) (Guthrie, 2004; Graham et al., 2016; Jakobsson et al., 2017). Concurrently, the warming climate and glacial recession restricted tundra habitat to higher latitudes as boreal forests expanded northward. *Sorex ugyunak*, isolated on the eastern side of the Bering strait, tracked the northern recession of tundra habitat, while western populations of *Sorex cinereus* (sensu stricto) expanded northward into Alaska with the advance of forest habitat.

In this chapter I make use of a multilocus genomic dataset of microsatellites, mitochondrial cytochrome b sequences, and ~11,000 single nucleotide polymorphisms (SNPs), to investigate the contributions of genetic drift and divergent selection to species differentiation in this system. I predict that as a result of the repeated and rapid geographic isolation of ancestral shrews in Beringia (on islands and separate continents), genetic drift has had a strong role in overall species differentiation. I also predict that due to small island size and complete isolation with no gene flow, *Sorex pribilofensis* in particular has experienced strong genetic drift, and that this has been the driving force in its genetic differentiation from mainland shrews. If true, this would have severe implications for future persistence of this species, given that significant fixation through the genome as a result of drift would limit its ability to exhibit an adaptive response to ongoing change. To test these predictions, I first analyze genome-wide SNP data to

detect loci under selection and investigate the contribution of selectively neutral versus non-neutral loci to species differentiation within this system (as opposed to explicitly investigating speciation based on all available loci, as reported in Chapter 2). I then compare the levels of genetic diversity of *Sorex pribilofensis* to sister taxa, estimate contemporary effective population sizes, and infer demographic changes over time across all included species. I then conduct linear regression analyses to test the prediction that genetic drift has led to the observed genetic differentiation in this system overall. Finally, I further test the prediction that *Sorex pribilofensis* has differentiated primarily through genetic drift by comparing its genetic differentiation from mainland shrews at the loci most impacted by drift to its differentiation at the loci least impacted by drift.

Methods

Sample collection and preparation of genetic markers

The individual shrews analyzed in this chapter are the same as those used in Chapter 2 (Table 2.1; Fig. 2.1; Table A1). The microsatellite, Cytb, and SNP datasets generated through methods detailed in Chapter 2 are further analyzed in this chapter to test predictions about the contributions of divergent selection and genetic drift in this species complex. After filtering each dataset for quality, the final microsatellite dataset included 17 loci for 79 individuals across 6 species, the Cytb dataset included 855-1140 bp sequences for 104 individuals across 6 species, and the SNP dataset included 11,989 polymorphic SNPs for 90 individuals across 4 species.

Detection of loci under selection

To conduct population genetic analyses and to investigate the roles of genetic drift and natural selection in this system, datasets of putatively neutral loci (hereafter: neutral dataset) and

loci presumably under selection (hereafter: non-neutral dataset) were required. To this end, I identified candidate loci under selection within the SNP dataset using two methods, Bayescan (Foll & Gaggiotti, 2008) and the R package *PCadapt* (Luu et al., 2017). Both Bayescan and *PCadapt* rely on population differentiation to identify candidate loci under selection. However, exploratory analyses showed virtually no population differentiation within *S. pribilofensis*. Additionally, I was unable to identify any loci under divergent selection in pairwise comparisons between *S. pribilofensis* and the other three species, possibly due to high background pairwise F_{ST} rates that could conceal even high differences in allele frequencies between species, the indicator of positive selection via this analysis. So in order to identify candidate loci within the three members of the Beringian clade, I treated *S. pribilofensis*, *S. portenkoi*, and *S. ugyunak* as three subunits of one larger evolutionary unit, considering that they are sister taxa and share a common ancestor based on genetic evidence from both mitochondrial and nuclear data. To identify candidate loci under selection between only these three species, *S. cinereus* (the outgroup taxon) was first removed from the dataset and SNPs that were monomorphic within the other three species were excluded, leaving 5,831 polymorphic SNPs for analysis of potential selection. Conservatively, only loci detected by both Bayescan and *PCadapt* were retained for the non-neutral dataset and only loci identified by neither method were retained for the neutral dataset.

Bayescan v2.1 was run with 20 pilot runs of 5,000 generations, followed by 100,000 generations and an additional burn-in of 50,000 iterations (for a total of 150,000 generations). The prior odds for the neutral model were set at 10 and candidate loci were identified using a false discover rate (FDR) of 0.1. Given I was primarily concerned with creating the neutral dataset, I set these values to err on the side of including false positives in the non-neutral dataset,

as opposed to setting conservative values which might leave true loci under selection in the neutral dataset. False positives in the non-neutral dataset were then mitigated by only including loci identified by both Bayescan and *PCadapt*.

PCadapt v4.3.3 detects outlier loci based on principal component analysis (PCA) by assuming that loci excessively related with population structure are candidates for local adaptation. After examining the scree plot and PCA plots, I kept only the first 2 principal components, as they were sufficient to describe the relevant structure between the three species (Fig. 3.1). Candidate loci under selection were identified after first transforming p-values to q-values and using a FDR of 0.05.

Contributions of selection and genetic drift to species structure

To investigate the relative roles of divergent selection and genetic drift in the evolution of these shrews, I characterized species structure separately for both the non-neutral and neutral SNP datasets. I had two predictions regarding the contribution of these evolutionary processes to the observed species structure in this system. If the main process leading to the observed species structure is divergent selection (adaptation of each species to their local environments), I expected to find distinct species structure in the non-neutral dataset and less distinct species structure in the neutral dataset. Alternatively, if genetic drift has played a major role in leading to the observed species structure, I expected to see distinct species structure in both the neutral and non-neutral datasets. Due to the inherent nature of the non-neutral dataset (comprised of the SNPs most significantly associated with species structure) I expected to observe structure between species based on these loci regardless of the evolutionary contribution of adaptive selection in this system. The main difference between these scenarios is the expected species structure within the neutral dataset. However, I acknowledge that distinct structure from both

datasets may indicate that both selection and drift have played significant roles. For both datasets, the species structure was characterized with DAPC and Structure, in all cases following the same steps as outlined previously (in Chapter 2) for the total SNP dataset.

Testing the role of genetic drift within this system of shrews

Genetic drift is strongest in populations with small effective population sizes, that have gone through bottlenecks, and/or exhibit founder effects (Charlesworth, 2009; Dobzhansky & Pavlovsky, 1957). One of the major consequences of strong genetic drift is the reduction of genetic variability at any given locus due to random fluctuations of allele frequencies (Wright, 1931). In the most extreme case these random fluctuations lead to the random fixation of one allele and the loss of all other alleles. Therefore, two main predictions stem from the hypothesis that strong genetic drift has led to the observed species differentiation in this system; (1) we would expect island shrew species to exhibit smaller effective population sizes and have less genetic variability than mainland shrew species, and (2) we would expect a negative relationship between estimates of intraspecific genetic variability (indices of genetic drift) and the degree of interspecific genetic differentiation (Chakraborty & Nei, 1977; Funk et al., 2016; Jordan & Snell, 2008).

To test the first prediction, I compared the contemporary effective population sizes and demographic histories of each shrew species represented in the SNP dataset. Contemporary effective population sizes (N_e) were estimated in NeEstimator v2.1 using the linkage disequilibrium method and assuming random mating (Do et al., 2014). As this method assumes that markers are selectively neutral, I estimated N_e using only the microsatellite and neutral SNP datasets. I estimated N_e separately for each species and quantified the uncertainty with jackknifed 95% confidence intervals.

Demographic histories for each species were inferred in Stairway Plot v2.0 (Liu & Fu, 2020). This method is based on the SNP frequency spectrum (SFS) and does not assume selectively neutral markers. Additionally, this method allows for folded SFS, thereby not requiring the ancestral alleles of SNPs and making it ideal for non-model organisms. To create the SFS for each species, the cleaned and trimmed reads generated by the *process_radtags* module in Stacks v2.54 were aligned to the *Sorex araneus* reference genome with the software package BWA (Li & Durbin, 2009). Maximum likelihood (ML) estimation of the folded SFS for each species was then computed in ANGSD v0.933 (Korneliussen et al., 2014) under a SAMtools model. To be processed all reads had to have a mapping quality (MapQ) score of 25 or higher. From there, all sites had to be present in at least 80% of individuals of the given species and have a base quality score of 13 or higher. These filters retained 835,377 sites for *S. pribilofensis*, 795,092 sites for *S. cinereus*, 837,805 sites for *S. ugyunak*, and 855,837 sites for *S. portenkoi*. The folded SFS for each species was used to infer changes in effective population sizes over time. For each species I assumed an average generation time of 1 year (Hope et al., 2010). In the absence of direct estimates of *Sorex* mutation rates for the nuclear genome I used the default setting of 1.2×10^{-8} per site per generation, a value similar to those estimated for other mammals (Campbell et al., 2012; Kumar & Subramanian, 2002; Liu et al., 2014; Roach et al., 2010; The 1000 Genomes Project Consortium, 2010) and used in other demographic history analyses of mammals with unknown mutation rates (Benazzo et al., 2017; Beynon et al., 2015; Hansen et al., 2018; MacLeod et al., 2013). Median effective population sizes and 95% confidence intervals were estimated based on 200 bootstrap replicate analyses.

To finish testing the first prediction, I compared measures of genetic variability for the microsatellite, Cytb, and total SNP datasets. For the microsatellite dataset, observed

heterozygosity (H_o), expected heterozygosity (H_e), allelic richness (Ar), and F_{IS} were estimated using the R packages *adegenet* (Jombart, 2008), *hierfstat* (Goudet, 2005), and *PopGenReport* (Adamack & Gruber, 2014). For the Cytb dataset, nucleotide diversity (π) and haplotype diversity (H_d) were estimated in DnaSP v6.12 (Rozas et al., 2017). For the SNP datasets, π and F_{IS} were estimated in Stacks v2.54 (Rochette et al., 2019), and H_o , H_e , and Ar were estimated with the R packages *adegenet* (Jombart, 2008), *hierfstat* (Goudet, 2005), and *dartR* (Gruber et al., 2018). While genetic drift may impact any locus, it is capable of acting strongest on loci evolving in the absence of or under weak selection (i.e. putatively neutral loci) (Ohta & Tachida, 1990; Ohta, 1992). I therefore also estimated the same measures of genetic variability (excluding F_{IS}) for the neutral SNP dataset as I did for the total SNP dataset.

To test my second prediction, that drift has resulted in differentiation through the loss of genetic variability, I analyzed the neutral SNP dataset and microsatellite dataset. For the neutral SNP dataset, I used three measures of genetic variability (H_o , H_e , Ar), estimated as described above. Similarly, I used three measures of genetic differentiation (Nei's F_{ST} , Nei's G_{ST} , Jost's D), which were estimated using the neutral dataset in the R packages *mmod* (Winter, 2012) and *hierfstat* (Goudet, 2005). This method of comparing differentiation and diversity assumes that intraspecific genetic variability is a reasonable index of the magnitude of historical drift, which should be true given that geographical isolation has restricted gene flow in this system.

Therefore, gene flow is not expected to contribute to intraspecific genetic variability. To remove the possibility of time since divergence contributing to the observed genetic differentiation between species, only pairwise estimates of genetic differentiation between each of the three Beringian shrews (*S. pribilofensis*, *S. ugyunak*, and *S. portenkoi*) and *S. cinereus* were used, as current and previous (Demboski & Cook, 2003; Hope et al., 2012) phylogenetic analyses have

shown *S. cinereus* to be an outgroup to the Beringian clade. Additionally, I have minimized potential influence of selection on observed genetic differentiation by using the neutral dataset. I then performed linear regression analyses in RStudio v1.3.1093 (RStudio Team, 2020), with each measure of interspecific genetic differentiation (F_{ST} , G_{ST} , Jost's D) as a response variable and each measure of intraspecific genetic variability (H_o , H_e , A_r) as an explanatory variable, for a total of nine separate regression analyses. To account for the multiple hypothesis testing inherent to this approach, I adjusted the p-values to control for false discovery using the BH method (Benjamini & Hochberg, 1995).

I repeated these analyses using the microsatellite dataset for the same three measures of genetic variability (H_o , H_e , A_r) and the same three measures of genetic differentiation (F_{ST} , G_{ST} , Jost's D), which were estimated using the R packages *hierfstat* (Goudet, 2005) and *mmod* (Winter, 2012). I again used only the pairwise estimates of genetic differentiation between each Beringian species and *S. cinereus* to control for time since divergence, and again adjusted the p-values as described above.

Testing the role of genetic drift in the speciation of Sorex pribilofensis

The two predictions described above relate to the role of genetic drift in this system of shrews overall. A third prediction stems from the hypothesis that genetic drift due to insular isolation was the primary driver of rapid divergence for *S. pribilofensis*. Namely, if genetic drift was the primary factor in its speciation, then I expect to see higher differentiation at loci most impacted by drift, and lower differentiation at loci least impacted drift. This would indicate that without the effects of drift, *S. pribilofensis* would be less differentiated. Based on current and previous (Demboski & Cook, 2003; Hope et al., 2012) phylogenetic analyses, *S. pribilofensis*, *S. ugyunak*, and *S. portenkoi* are known to have differentiated from a common ancestral form, with

S. pribilofensis being the first to diverge. Given that any potentially non-neutrally evolving SNP was removed to create the neutral SNP dataset, it is a fair assumption that allele frequencies of the neutral SNPs have mostly been influenced by drift in the absence of selection. In the absence of strong selection and gene flow, these neutral SNPs should all experience genetic drift at a relatively equal rate (Ohta & Tachida, 1990; Wright, 1931). However while this rate should be equal within each species, it will vary between species, because the strength of genetic drift is related to the effective population size (Charlesworth, 2009). Given enough time and the absence of other evolutionary forces, the result of drift is the fixation of alleles. However, at any given timepoint, some alleles will remain unfixed through random chance. So, at a given timepoint, the most extreme result of drift is the random fixation of one allele and the loss of all other alleles. However, due to the random nature of genetic drift, allele frequencies for each SNP should be largely independent, meaning that while some alleles may have drifted to fixation, some, through chance, will have retained variability. Therefore, the differentiation of *S. pribilofensis* should primarily reflect the loss of alleles due to the random effects of genetic drift acting on shared genetic variation across the three species. Specifically, we would expect higher interspecific genetic differentiation at loci that are fixed for one allele (monomorphic) within *S. pribilofensis* than at loci that are still variable (polymorphic) within *S. pribilofensis*.

To test this third prediction, I first filtered the neutral SNP dataset to include only SNPs monomorphic within *S. pribilofensis*, and estimated two measures (G_{ST} , Jost's D) of interspecific (*S. pribilofensis* vs *S. portenkoi*; *S. pribilofensis* vs *S. ugyunak*) genetic differentiation at these SNPs. I then filtered the neutral SNP dataset to include only SNPs polymorphic within *S. pribilofensis*, and estimated the same measures of interspecific pairwise genetic differentiation at those SNPs. In order to ensure that monomorphic SNPs within *S. pribilofensis* reflected fixation

since island isolation and not fixation within an ancestral taxon prior to island isolation, I verified that all monomorphic SNPs within *S. pribilofensis* were still polymorphic across *S. portenkoi* and *S. ugyunak*. Similarly, I measured genetic variability within each species for both kinds of SNP to verify that genetic variability was not lower within *S. portenkoi* and *S. ugyunak* at SNPs monomorphic within *S. pribilofensis* than at SNPs polymorphic within *S. pribilofensis*.

To measure the uncertainty around these point estimates of genetic differentiation, I created 200 bootstrap samples of the neutral monomorphic SNP dataset and 200 bootstrap samples of the neutral polymorphic SNP dataset using the *chao_bootstrap* function in the R package *mmod* (Winter, 2012). This function creates bootstrap samples by resampling each defined group according to its population size. I then measured pairwise interspecific genetic differentiation for each bootstrap sample and tested the distribution of each set of interspecific genetic differentiation values for normality. The bootstrapping function tends to have a slight upward bias, so I centered the distributions around the point estimate by subtracting the difference between the mean of each sample and the point estimate from each interspecific genetic differentiation estimate (as suggested by the author of *mmod*). As all samples had normal distributions, I then conducted t-tests to test whether pairwise interspecific genetic differentiation was different between SNPs monomorphic within *S. pribilofensis* and SNPs polymorphic within *S. pribilofensis*.

Results

Detection of loci under selection

Using the 5,831 polymorphic SNPs between *S. pribilofensis*, *S. portenkoi*, and *S. ugyunak*, Bayescan detected 15 outlier loci while *PCadapt* detected 568 outlier loci (Figs. 3.1, 3.2). All 15 outlier loci detected by Bayescan were also detected by *PCadapt*. The 15 outlier loci

detected by both methods were used to create the non-neutral dataset, while the 5,263 loci that were not detected as outliers by either method were used to create the neutral dataset.

Contributions of selection and genetic drift to species structure

The optimal number of PCs to retain for the Discriminate Analysis of Principle Components (DAPC) was 3 for both the neutral and non-neutral datasets. The neutral dataset DAPC (Fig. 3.3) showed similar species relationships to those found using the total SNP dataset (Fig. 2.5). When projecting into the first and second PCs, *S. portenkoi* and *S. ugyunak* are closely aligned, while *S. cinereus* and *S. pribilofensis* are both independently isolated. The third PC provides more distinction between *S. portenkoi* and *S. ugyunak*. For the non-neutral dataset, all four species are spatially differentiated using only the first two principle components (Fig. 3.4). The most distant species based on the non-neutral dataset is *S. portenkoi*, as opposed to *S. cinereus* based on the neutral dataset.

Relationships similar to those observed in the DAPC analyses were evident in clustering analysis using Structure (Figs. 3.5, 3.6). For both the neutral and non-neutral SNP datasets, *a priori* species designations did not impact how individuals clustered or the most likely number of clusters, so here I report results for analyses without *a priori* species designations. The most likely number of genetic clusters was K=3 for both the neutral and non-neutral SNPs, but K=4 also had a high likelihood for the neutral dataset. Similar to Structure analysis for the total SNP dataset (Fig. 2.7), for the neutral SNP dataset when K=3 *S. cinereus* and *S. pribilofensis* form distinct genetic clusters while *S. portenkoi* and *S. ugyunak* group together. However, for the neutral SNP dataset *S. portenkoi* and *S. ugyunak* share some genetic similarity with *S. pribilofensis*, unlike for the total SNP dataset. This pattern is also seen for K=4 for the neutral SNPs, as the four species group independently while *S. ugyunak* and some *S. portenkoi*

individuals share genetic similarity with *S. pribilofensis*. Different clusters are observed in Structure analysis of the non-neutral SNP dataset, as *S. portenkoi* and *S. ugyunak* form distinct clusters while *S. cinereus* and *S. pribilofensis* group together for $K=3$. Neither $K=4$ or $K=5$ had high likelihood, and neither was able to recover the four species clusters for the non-neutral SNP dataset.

Testing the role of genetic drift in this system of shrews

Demographic analyses – Using the linkage disequilibrium method in NeEstimator v.21, estimates of effective population size (N_e) for both the microsatellite and neutral SNP datasets were projected to infinity for most species, likely due to small sample sizes (Tables 3.2, 3.3). Estimates of $N_e=\infty$ can be interpreted as the sample size being too small relative to the true effective population size for accurate estimation (Marandel et al., 2019; Nunziata & Weisrock, 2018). As many shrew species have very large census sizes (Dokuchaev, 1989; Whitaker, 2004), the sample sizes used for N_e estimation with the linkage disequilibrium method are likely too small in relation to provide reliable estimates. Given a higher sample size for *S. pribilofensis*, mean N_e estimates are potentially more accurate than for the other species (microsatellite estimate: 18.4; neutral SNP estimate: 2106.5), however the upper bound on the 95% confidence interval for both estimates is infinity.

Demographic histories inferred in Stairway Plot v2.0 with the neutral SNP dataset provided higher resolution for contemporary effective population sizes and for change in N_e through time (Figs. 3.7-3.11). The estimated contemporary N_e for *S. pribilofensis* was 43.33 (95%CI: 17.66, 928.26), the smallest of the four species represented in the neutral SNP dataset. Additionally, the effective population size of *S. pribilofensis* has been steadily shrinking since ~2 kya, which was the inferential limit for this species given the applied mutation rate and

generation time. The demographic histories of *S. cinereus*, *S. portenkoi*, and *S. ugyunak* could be inferred back to ~20 kya, around the time of the last glacial maximum (LGM). *S. ugyunak* had the largest estimated contemporary effective population size ($N_e=67,325$; 95% CI: 10,418, 203,365), with a demographic history of slow population decline followed by rapid growth around 5 kya. The relatively small population sample of *S. cinereus* from one general region of Alaska also had a large estimated contemporary effective population size ($N_e=21,661$; 95% CI: 407.87, 112,217), but with a demographic history of steady population decline followed by possible recent growth. Aside from *S. pribilofensis*, *S. portenkoi* had the smallest estimated contemporary effective population size ($N_e=2,879$; 95% CI: 836.23, 40,706), with a demographic history of steady population decline since ~8 kya. Recent estimates of N_e for all species except *S. pribilofensis* were associated with broad 95% confidence intervals.

Genetic variability – I found extremely low genetic variability across all datasets within *S. pribilofensis*; compared to sister taxa, most measurements of genetic variability were at least an order of magnitude lower (Tables 3.4-3.6). The highest genetic variability was recovered for *S. cinereus*, except at the microsatellite loci, where it was comparable with *S. portenkoi* and *S. ugyunak*. Genetic variability within the other island endemic species, *S. jacksoni*, was intermediate between *S. pribilofensis* and the mainland species. The inbreeding coefficient, F_{IS} , was estimated to be low for all taxa based on both the microsatellite and SNP datasets. The highest F_{IS} values were observed with the microsatellite dataset for *S. pribilofensis* (0.195) closely followed by *S. camtschatica* (0.186). Trends in genetic variability were similar for the total SNP dataset as for the neutral SNP dataset. Genetic variability estimates based on neutral SNPs are only reported for *S. pribilofensis*, *S. portenkoi*, and *S. ugyunak*, as they were the species

used to isolate the putatively neutral loci. Genetic variability estimates for these species were higher based on the neutral SNP dataset than the total SNP dataset.

Linear regressions – All linear regressions, for both the microsatellite and neutral SNP datasets, showed a negative relationship between genetic variability and genetic differentiation (Figs. 3.12, 3.13). The relationship was significant ($p < 0.05$) for all of the regressions using the microsatellite dataset and for three of the regressions using the neutral SNP dataset, although all regressions using the neutral SNP dataset had R^2 values of 0.96 or higher (Table B1). Generally, the relationship between genetic variability and genetic differentiation was most significant for the neutral SNP dataset with H_e as the measure of genetic variability, and for the microsatellite dataset with H_e and A_r as the measures of genetic variability. The least significant relationships for both datasets were based on Jost's D as the measure of genetic differentiation, although these relationships were still significant for the microsatellite dataset. In all cases, the species with the least genetic variability (*S. pribilofensis*) was the most differentiated from *S. cinereus*, while the species with the most genetic variability (*S. ugyunak* or *S. portenkoi*) were among the least differentiated. In comparison to the other species, *S. jacksoni*, which is endemic to a larger island (St. Lawrence) that has been isolated from the mainland for less time than St. Paul Island, exhibited intermediate genetic variability and differentiation from *S. cinereus*.

Testing the role of genetic drift in the speciation of Sorex pribilofensis

Of the 5,263 SNPs in the neutral dataset, 232 were polymorphic within *S. pribilofensis* and 5,031 were monomorphic (Fig. 3.14). All of the 5,031 monomorphic SNPs within *S. pribilofensis* were still polymorphic across *S. portenkoi* and *S. ugyunak*, and genetic variability measures were higher for *S. portenkoi* and *S. ugyunak* at monomorphic *S. pribilofensis* SNPs than at polymorphic *S. pribilofensis* SNPs (Fig. 3.14; Table 3.7). At its 232 polymorphic SNPs,

genetic variability measures for *S. pribilofensis* were similar to those of *S. portenkoi* and *S. ugyunak*. Each pairwise genetic differentiation measure between *S. pribilofensis* and each of *S. portenkoi* and *S. ugyunak* was significantly lower ($p < 0.001$) at polymorphic *S. pribilofensis* SNPs than at monomorphic *S. pribilofensis* SNPs (Fig. 3.15). At both monomorphic and polymorphic SNPs, *S. pribilofensis* was generally less genetically differentiated from *S. ugyunak* than from *S. portenkoi*. However, the difference in genetic differentiation at polymorphic and monomorphic *S. pribilofensis* SNPs was greatest in comparison to *S. portenkoi* (difference in G_{ST} of 0.110; difference in Jost's D of 0.065).

Discussion

My analysis of a tiered genomic dataset, comprised of locus-sets with different levels of resolution through mode of inheritance, rate of evolution, and their inferred neutrality (or lack thereof), revealed that genetic drift has been a predominant force in speciation within this complex of closely related shrews. This result is supported by four independent lines of evidence. First, genetic clustering analyses show significant genetic clusters at neutrally evolving SNPs. Second, genetic variability and effective population sizes are lower for island species, especially for *S. pribilofensis*, in comparison to mainland species. Third, there is a significant negative relationship between genetic variability and genetic differentiation, showing that species with less genetic variation are more highly differentiated. Finally, *S. pribilofensis* is more differentiated from its closest mainland relatives (*S. portenkoi* and *S. ugyunak*) at the loci that have been most impacted by genetic drift within *S. pribilofensis* (monomorphic loci), and less differentiated at the loci least impacted by genetic drift (polymorphic loci). While this doesn't preclude the potential for divergent selection to have caused some degree of differentiation, it is

clear that speciation in this system, especially for *S. pribilofensis*, has been heavily influenced by genetic drift. This adds to a growing body of literature showing that genetic drift can play a leading role in the divergence of isolated taxa (Funk et al., 2016; Jordan & Snell, 2008; Prentice et al., 2017; Toczydlowski & Waller, 2019), and provides evidence that speciation through drift, under the right historical biogeographic conditions, can occur on a remarkably short evolutionary timescale.

Species structure at neutral vs non-neutral SNPs

Species clustering analyses for the neutral SNPs showed similar relationships as observed using the total SNP dataset (Figs. 2.4-2.7). DAPC using the neutral SNPs and the first two principle components shows significant genetic structure, with *S. cinereus* and *S. pribilofensis* grouping independently, while *S. ugyunak* and *S. portenkoi* group closely yet distinctly. While these first two principle components mostly describe the genetic differentiation of *S. cinereus* and *S. pribilofensis*, the third principle component allows for more discrimination between *S. portenkoi* and *S. ugyunak*, showing that while the genetic differentiation between these species is not as great as that observed for *S. cinereus* and *S. pribilofensis*, there is still relevant genetic structure between them. This trend is also observed in the results of the Structure analysis. As for the total SNP dataset, the most likely number of genetic clusters is K=3, while K=4 also shows relevant structure and has a high likelihood. Using the neutral SNPs, *S. cinereus* and *S. pribilofensis* again group distinctly, while *S. portenkoi* and *S. ugyunak* are grouped together for K=3. Perhaps most interestingly, this genetic structuring mirrors that observed with the total SNP dataset. Since the neutral SNP dataset excludes all loci showing signals of potentially being under selection, this suggests that neutral evolutionary processes, such as genetic drift, are strongly contributing to the observed genetic structure within this system.

In contrast to structuring observed with the total SNP dataset, it appears that for neutral SNPs, *S. ugyunak*, and to a certain extent *S. portenkoi*, share a non-trivial amount of genetic similarity with *S. pribilofensis*. It is possible that the ability to completely distinguish between *S. ugyunak* and *S. portenkoi* is hindered without the inclusion of the non-neutral SNPs, and that this partial grouping with *S. pribilofensis* mostly reflects increased similarity between *S. ugyunak* and *S. portenkoi*. This rationale is supported by the results of genetic clustering analyses using the non-neutral dataset. Initial attempts were made to identify loci under selection within *S. pribilofensis* only, but a complete lack of intra-specific population structuring precluded detection of selection, given the genetic differentiation-based approach of Bayescan and *PCadapt*. Attempts to identify loci under selection in pairwise comparisons of *S. pribilofensis* and each of *S. portenkoi* and *S. ugyunak* also returned no results, possibility due to high background rates of genetic differentiation between these species. Given that no loci under selection were detected specifically for *S. pribilofensis*, it is likely that the putatively non-neutral SNPs detected between *S. pribilofensis*, *S. portenkoi*, and *S. ugyunak* largely reflect non-neutral evolutionary processes, such as divergent selection, between the latter two mainland species. DAPC and Structure analysis using the 15 non-neutral SNPs indeed show that the largest species differentiation at these SNPs is between *S. portenkoi* and *S. ugyunak*. In the DAPC plot, these two species are the most distant of any pairwise comparison, while *S. cinereus* and *S. pribilofensis* group closely and near the origin of the plot. This central positioning suggests that the principle components describing the genetic variation at these non-neutral SNPs provide little information on the differentiation *S. cinereus* and *S. pribilofensis*. This makes sense given *S. cinereus* was not included in the identification of non-neutral SNPs, and shows that these non-neutral SNPs likely are not important for the evolution of *S. pribilofensis* either. Instead, they

may well reflect adaptive divergence between *S. portenkoi* and *S. ugyunak*. Given that these species have larger effective population sizes than *S. pribilofensis*, selection would indeed be expected to act more effectively. This conclusion is supported by the genetic clusters observed through Structure analysis, as $K=3$ was the most likely number of genetic clusters, with *S. cinereus* and *S. pribilofensis* grouping together, and no relevant species structure observed beyond this number of groups. However, it is worth noting that *S. portenkoi* and *S. ugyunak* are still differentiated based on neutral SNPs, making it unlikely that their divergence is due entirely to non-neutral evolutionary processes. Overall, this analysis suggests that there may be some non-neutral differentiation between *S. portenkoi* and *S. ugyunak*, but that the observed differentiation for all four species is best described by neutrally-evolving SNPs.

Demographic histories

Estimates of effective population sizes using the linkage disequilibrium method in NeEstimator were largely uninformative, as N_e for most species was projected to infinity. This is likely due to small sizes in relation to the true effective population size; for example, in one estimation of the sample size needed for accurate estimation of N_e , it was determined that the sample size should be at least 1% of the census size (Marandel et al., 2019). Even given a low estimation of total population sizes of around ~10,000 individuals per shrew species, this would require sample sizes of at least 100 for each species. I was able to provide point estimations of N_e for *S. pribilofensis* using both the microsatellite and neutral SNP datasets, although the upper limit of the 95% confidence interval for these estimations was still projected to infinity. This could simply be a reflection of the larger sample size for *S. pribilofensis*; with increased sampling of the other species, N_e estimation may become more feasible. However, the fact that a point estimate was possible for this species lends further support for a truly smaller N_e of *S.*

pribilofensis. The other species for which a non-infinite estimate of N_e was obtained was *S. portenkoi*. Given a limited sample size and only two sampling localities, it's possible the low estimated N_e for this species reflects a local population estimate as opposed to for the entire species. Evidently, additional samples would help to refine inference from this analysis.

Demographic histories inferred with Stairway Plot 2 provide a glimpse into the relative changes in effective population size of each species over time, in addition to an estimation of contemporary effective population sizes. With the exception of *S. pribilofensis*, contemporary effective populations sizes for these species are associated with very large 95% confidence intervals, making it difficult to accurately infer changes in population size through time for *S. cinereus*, *S. ugyunak*, and *S. portenkoi*. However, there is strong evidence that *S. pribilofensis* has a significantly lower effective population in comparison to *S. cinereus* and *S. ugyunak*, based on the lack of overlap of 95% confidence intervals. While the contemporary difference in N_e is not significant between *S. pribilofensis* and *S. portenkoi*, the small recent N_e of *S. portenkoi* could again be reflective of limited sampling. The demographic history of *S. pribilofensis* indicates a steady decline in N_e since at least 2 kya, well within the history of island isolation beginning ~14 kya. It is likely that the true population size has been declining ever since island isolation, as the island continuously shrunk until it reached its current size around 6 kya, but given the limited genomic variability of *S. pribilofensis* the inferential limit for this analysis was around 2 kya. In addition to a decline in the true population size, this decline in N_e likely also reflects a history of erosion of genomic diversity through repeated and possibly sustained census population fluctuations and bottlenecks.

As for the other species, *S. cinereus* shows a history of steady population decline since the LGM ~20 kya. Given that only Alaskan *S. cinereus* were included in this analysis,

representing individuals from only one region of a species that occurs across much of North America, these results shouldn't be interpreted as representative of range-wide demographic trends, but instead viewed in light of the evolutionary history of Alaskan *S. cinereus*. Previous work has shown that *S. cinereus* likely did not occur in the Beringian refugium formed during the last glacial period (~75-11 kya), but instead tracked the northern expansion of boreal forests since the LGM (Hope et al., 2012). As such, populations at the leading edge of this range expansion would have experienced continuously elevated isolation from populations nearer the center of the species' range, resulting in a slowly shrinking effective population size (Excoffier et al., 2009; Nichols & Hewitt, 1994). Meanwhile, *S. ugyunak* shows a demographic history of slight decline since 20 kya followed by rapid expansion ~5 kya. This population growth could be reflective of an increase in its preferred mesic tundra habitat as glaciers slowly receded following the LGM (Hope & Elias, 2019).

Genetic drift in this system of shrews

If genetic drift is driving differentiation in this system of shrews, my first prediction was that island species would have lower effective population sizes and less genetic variability in comparison to mainland species. Despite the difficulties in obtaining accurate estimations of effective population size and evaluating their significance across species, there does appear to be evidence for a smaller, and declining, effective population size of *S. pribilofensis* compared to mainland species. Additionally, both island species (*S. pribilofensis* and *S. jacksoni*) have reduced genetic variability in comparison to the mainland species. In the case of *S. pribilofensis*, most estimates of genetic variability were at least an order of magnitude smaller than the observed genetic variability of the mainland species. As an endemic to a larger island (St. Lawrence) than *S. pribilofensis*, *S. jacksoni* exhibits more genetic variability than *S.*

pribilofensis, but still less than the mainland species. It is worth noting that estimations of genetic variability were higher for each of *S. pribilofensis*, *S. portenkoi*, and *S. ugyunak* when using the neutral SNP dataset in comparison to the total SNP dataset. This is likely a by-product of the methods used to create the neutral and non-neutral SNP datasets, because in order to identify neutral and non-neutral SNPs between these three species, *S. cinereus* was removed and resulting monomorphic loci filtered out prior to analysis. This removal of invariant sites would inherently lead to an increase in observed estimations of genetic variation, so there is likely no meaningful biological reason for the difference in these estimations. Taken together, the results of effective population size and genetic variability analyses point to genetic drift as a strong evolutionary force in the differentiation of island shrew species, evidenced by smaller effective population sizes and reduced genetic variability (Charlesworth, 2009; Wright, 1931).

My second prediction was that interspecific genetic differentiation would be significantly associated with reduced intraspecific genetic variation. Results from linear regressions based on both microsatellite and neutral SNPs indicate a strong negative relationship between all measures of genetic variability and all measures of genetic differentiation, and all relationships were statistically significant for the microsatellites and three out of nine relationships were for the neutral SNPs. The six non-significant relationships for the neutral SNPs still show a strong negative relationship, indicated by high R^2 values, but with only three data points is difficult for these regressions to reach statistical significance. By using *S. cinereus* as the outgroup species for each pairwise measure of genetic differentiation, these measures are essentially standardized for time since divergence. Therefore, any observed difference in pairwise differentiation must be due to an evolutionary process, not a longer period of isolation. Additionally, both the microsatellites and neutral SNPs represent neutrally evolving genetic markers, minimizing the

likelihood that non-neutral evolutionary processes like divergent selection have influenced these results. Reduced genetic diversity is a known consequence of genetic drift, so the negative relationship between genetic variability and genetic differentiation strongly indicates that drift has been a driving force of differentiation in this system. This kind of analysis has proven effective in tying genetic drift to differentiation in other systems (Funk et al., 2016; Jordan & Snell, 2008; Prentice et al., 2017; Whiteley et al., 2010).

Genetic drift and speciation in Sorex pribilofensis

My third prediction related specifically to the speciation of *S. pribilofensis*. Due to a smaller effective population size, genetic drift should have had a stronger influence on allele frequencies for *S. pribilofensis* than for *S. portenkoi* and *S. ugyunak* (Charlesworth, 2009). In order to show that this higher rate of genetic drift was a primary factor in the speciation of *S. pribilofensis*, I tested its differentiation from *S. portenkoi* and *S. ugyunak* at the loci most impacted by drift versus at those least impacted. To do this, I divided the neutral SNP dataset as such, using SNPs monomorphic within *S. pribilofensis* (fixed for one allele) as the loci most impacted by drift, and SNPs polymorphic within *S. pribilofensis* (retaining two alleles, at any frequency) as the loci least impacted by drift. I then showed that *S. pribilofensis* was significantly more differentiated from *S. portenkoi* and *S. ugyunak* (two of its most closely related sister species) at its monomorphic SNPs than at its polymorphic SNPs.

Without the influences of selection or gene flow, all the neutral SNPs within *S. pribilofensis* should be experiencing drift at an equal rate, but due to the random nature of drift acting on initial allele frequencies, some of these loci became monomorphic while some remained polymorphic (Ohta, 1992; Wright, 1931). The finding of less differentiation at neutral SNPs that are polymorphic within *S. pribilofensis* suggests that these loci are reflective of the

genetic variability of the most recent common ancestor (MRCA) of *S. pribilofensis*, *S. portenkoi*, and *S. ugyunak*. Through random chance, these loci have deviated less from their starting frequencies within *S. pribilofensis* after island isolation, resulting in *S. pribilofensis* being less differentiated from *S. portenkoi* and *S. ugyunak* at these loci. The finding of higher differentiation at neutral SNPs that are monomorphic within *S. pribilofensis* suggests that alleles at these loci drifted to fixation after island isolation, while they remained variable within *S. portenkoi* and *S. ugyunak*. It is unlikely that these alleles were fixed in the MRCA because they remain variable across *S. portenkoi* and *S. ugyunak*. If these alleles were all fixed prior to isolation of *S. pribilofensis*, then mutations would have had to subsequently arise and persist at each locus to account for the variation across *S. portenkoi* and *S. ugyunak*. The most parsimonious answer is that strong genetic drift, acting on genetic variation initially shared by the three species, led to the fixation of these alleles in *S. pribilofensis*, while genetic drift was not strong enough for alleles at these loci to become fixed across *S. portenkoi* and *S. ugyunak*. These results further support genetic drift driving very rapid speciation of *S. pribilofensis*.

Conservation implications

This work has several implications for the conservation and management of *S. pribilofensis*. The findings of limited genetic diversity and a small effective population size in relation to mainland species suggest that this species is vulnerable to progressive and stochastic environmental change through a reduction in overall fitness and adaptive potential (Hohenlohe et al., 2021; Reed & Frankham, 2003). Without the ability to disperse to other suitable environments, future responses of *S. pribilofensis* to ongoing change will have to be through plasticity or adaptation (Colella et al., 2020). As such, the management action most likely to be successful for the conservation of *S. pribilofensis* is the maintenance of suitable environmental

conditions on Saint Paul Island. This includes preservation of existing habitat, detection and eradication of any introduced species (e.g. rats, non-native flora, pathogens, etc.), and maintenance of favorable climatic conditions. The preferred habitat of *S. pribilofensis* has been previously described by Byrd & Norvell (1993); in short, they found the species to be most abundant in tall-plant vegetation, specifically in dune and grass-umbel habitats, which covers 54.9% of the island. As for the threat of invasive species, there is already evidence that *S. pribilofensis* is infected by a novel parasite with unknown pathogenicity (Hope et al., 2016) and preliminary analysis suggests the recent introduction of non-native slugs on St. Paul Island (Wiens, unpublished data). The effects of these introductions on *S. pribilofensis* are unknown, and work is needed to understand the complex host-parasite dynamics at play, including the potential role of new intermediate hosts in activating complex parasite life-cycles coupled with the process of ecological fitting under novel environments (Combe et al., 2021). Finally, even if the strongest possible action is taken to curb anthropogenic climate change, global climate will continue to change for the foreseeable future (Collins et al., 2013). Therefore, the primary management priorities should be the preservation of preferred island habitats and the control of invasive species.

While *S. pribilofensis* exhibits extremely low genomic diversity, it does not display a strong signal of inbreeding, as indicated by the inbreeding coefficient, F_{IS} , for either the microsatellite ($F_{IS}=0.195$) or SNP ($F_{IS}=-0.0014$) datasets. The implication is that while the overall genetic diversity of *S. pribilofensis* remains low, there is enough random mating to allow remaining genomic variability to persist (Hohenlohe et al., 2021; Keller, 2002). In other words, there is enough random mating to prevent high genomic linkage, so inheritance of one genotype doesn't strongly predict the inheritance of another. Likewise, while the observed

decline in effective population size over time is likely reflective at least in part of decline in census population size, it also reflects the degree of genetic similarity (i.e. limited genetic variability) within *S. pribilofensis*. The effective population size is usually smaller than the true population size, because it describes “the size of an ideal, panmictic population that would experience the same loss of genetic variation, through genetic drift, as the observed population” (Hohenlohe et al., 2021). This line of reasoning is supported by observational evidence for the census size of this shrew by three expeditions to St. Paul island in the 1900s, which all found it to be abundant on the island (Byrd & Norvell, 1993; Fay & Sease, 1985; Jackson, 1928). These observations are concurrent with the high success of trapping efforts carried out for this thesis, as 27 shrews were trapped in 288 trap nights. This anecdotal evidence further suggests that the census size of this shrew is capable of vastly exceeding the effective population size. Overall, these interpretations suggest that the single population of *S. pribilofensis* is currently locally common despite small total range size and incredibly limited genetic diversity. However, this may be subject to change quickly if/when progressive and stochastic environmental changes occur on Saint Paul Island. Therefore, the second management priority of *S. pribilofensis* should be regular population monitoring in order to detect any potential decrease in population sizes or increase in levels of inbreeding. Importantly, given the very high genetic distinction of this shrew from its sibling species as well as high ecological differences between respective habitats, any efforts at genetic rescue from other shrew species should be considered only as a last resort and after extensive field survey efforts that definitively (through scientific data) indicate a high risk of imminent extinction.

Finally, sustainable sampling and archival of additional specimens in natural history museums will enable continued research on the evolution of *S. pribilofensis*, which will enhance

conservation efforts and promises to increase our understanding of species-specific and general evolutionary responses to insular endemism. For example, to my knowledge only relatively small areas of the island have been sampled for *S. pribilofensis*. Determining the extent to which *S. pribilofensis* occupies other localities on Saint Paul Island, as well as any of the neighboring islands where it has never been detected (St. George Island 40 mi. south, and Walrus/Otter Islands, small satellite islands to St. Paul), and which habitats it prefers will not only be beneficial to conservation efforts, but provide a clearer picture of any potential genetic structure between populations. Another conservation action might be to establish additional populations of shrews on satellite islands where no other native mammal species are present, as this has been successful elsewhere for birds (Miskelly & Powlesland, 2013). However, this should only take place as a result of careful planning, as each island in this region represents a fragile ecosystem coupled with critical bird breeding habitat.

Importantly, sustainable and holistic sampling, preparation and preservation of specimens in museums has minimal negative side effects on natural population densities, and provides for spatial and temporal documentation of biodiversity while allowing for additional research avenues to be pursued in the future (Hope et al., 2018). Voucher specimens store a wealth of information and are critically important for an understanding of broad biological processes, as they provide documentation of associated biodiversity (e.g. viruses, bacteria, helminth parasites, etc.), the biological/chemical composition of diet, and a genetic and morphological record of populations/species (McLean et al., 2016). Sustainable sampling also provides a time series of specimens, the utility of which is becoming increasingly realized, as they allow for the direct investigation of molecular to community-level responses to environmental change through time (Bi et al., 2019; Moritz et al., 2008). Ongoing sampling efforts will also provide for more

accurate estimations of the inbreeding coefficient and effective population size change. An especially promising line of study is analysis of whole genomes, which is becoming increasingly tractable for non-model organisms (Ellegren, 2014; Jones & Good, 2016). This allows for the investigation of functional genetic diversity, and could provide insight into the potential ability of island species to retain variability at genes that code for important traits, despite virtual genomic flat-lining (Robinson et al., 2016). This would also shed light on the role of adaptive divergence in this system, and increase our understanding of the interaction between natural selection and genetic drift in the divergence of island taxa in general.

Genetic drift in the Arctic

Biogeographical conditions in this system have predisposed biodiversity, and specifically *S. pribilofensis*, to strong genetic drift. Initial expansion of an ancestral shrew taxon into Beringia during the Sangamon and subsequent isolation of this taxon during the Wisconsinan likely eroded initial diversity within Beringian species of the *cinereus* complex through pioneer expansion (Hewitt, 2004). Since the LGM (~20 kya), the range of this widespread ancestral taxon was fragmented, and populations were isolated on islands in the Bering Sea and on either side of the Bering Strait. The relatively simple community and habitat structure of this Arctic system is coupled with harsh conditions, especially on the islands, which experience even more depauperate community and habitat structure through faunal/floral relaxation. A lack of heterogeneity in local environment and a decrease in community interactions likely diminished the selective pressures at play in the system (Miller et al., 2019). For *S. pribilofensis*, small island isolation cut off gene flow with sister taxa, and led to a decreased effective population size, through some combination of founder effects, bottlenecks, and stochastic fluctuations in population size. Of consequence, this island isolation began only ~14 kya when St. Paul island

formed, meaning the consequences of strong genetic drift were able to accumulate on a short timescale. This rapid observed genetic differentiation was likely magnified by rapid turnover through short generation time.

By taking account of the biogeographic conditions that have led to speciation through genetic drift in this system, a predictive framework can be implemented for the role of drift in other high-latitude systems. Importantly, biogeographic histories that result in small effective populations sizes increase the potential for strong genetic drift. This may occur within populations of leading-edge range expansions responding to extreme environmental gradients, or populations which have become highly isolated, both on true islands and through habitat fragmentation across a previously connected landscape. Additionally, divergence through drift can be expected to happen more rapidly for species with shorter generation times, which as I have shown can lead to remarkably rapid speciation. While fluctuations in allele frequencies due to genetic drift are random, the outcome of genetic drift – differentiation through loss of diversity – is not, and can be expected within other taxa. Given similar biogeographic histories and environmental conditions, the contribution of drift to differentiation in other Arctic systems is an evolutionary outcome with pertinent implications for the distribution of biodiversity, and which is still under-explored.

Conclusions

This study shows that genetic drift has led to the rapid speciation of an island endemic, and adds to a growing body of literature documenting the contribution of genetic drift to the differentiation of isolated taxa (Funk et al., 2016; Jordan & Snell, 2008; Prentice et al., 2017; Toczydlowski & Waller, 2019). The ability of genetic drift to cause rapid divergence on such a short evolutionary timescale has major implications for this species and others as the

Anthropocene progresses. As island endemics are continuously forced to react to accelerating environmental change, an understanding of the legacy of genetic drift will become increasingly relevant for the conservation and persistence of biodiversity.

Figures

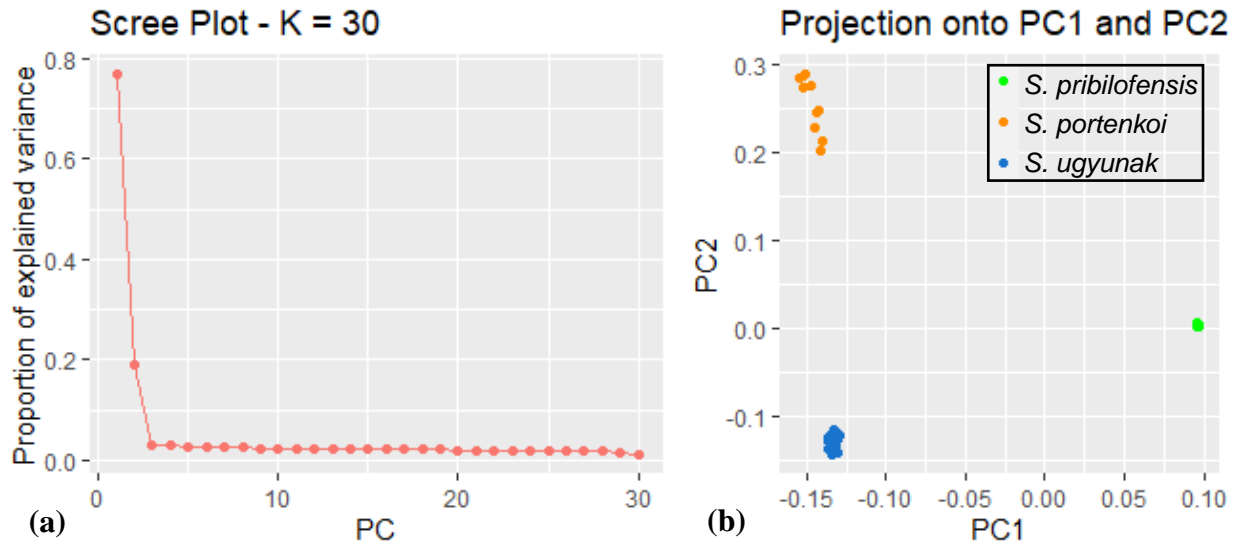


Figure 3.1 Detection of outliers with *PCadapt*

The population structure, identified through principle component analysis, used to identify SNPs under selection by *PCadapt*. (a) The proportion of variance explained by the first 30 principle components in the PCA. (b) Only the first two PCs were kept to identify SNPs under selection, as they were sufficient to describe the relevant structure between the three species.

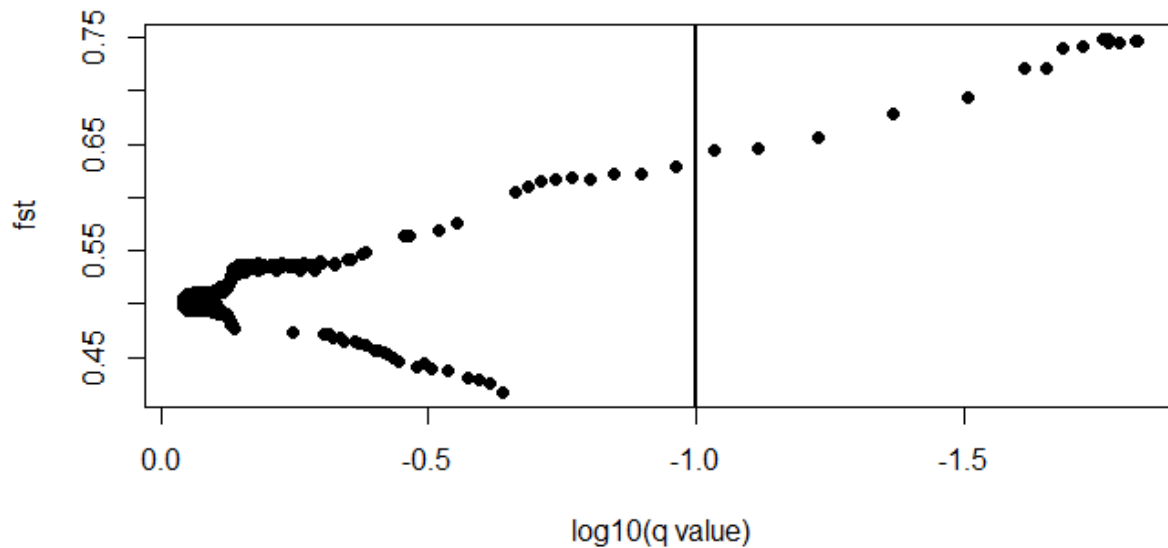


Figure 3.2 Detection of outliers with Bayescan

Outlier SNPs detected by Bayescan with a FDR of 0.1. Dots to the right of the solid black line indicate SNPs significantly associated with population structure, as indicated by higher F_{ST} values.

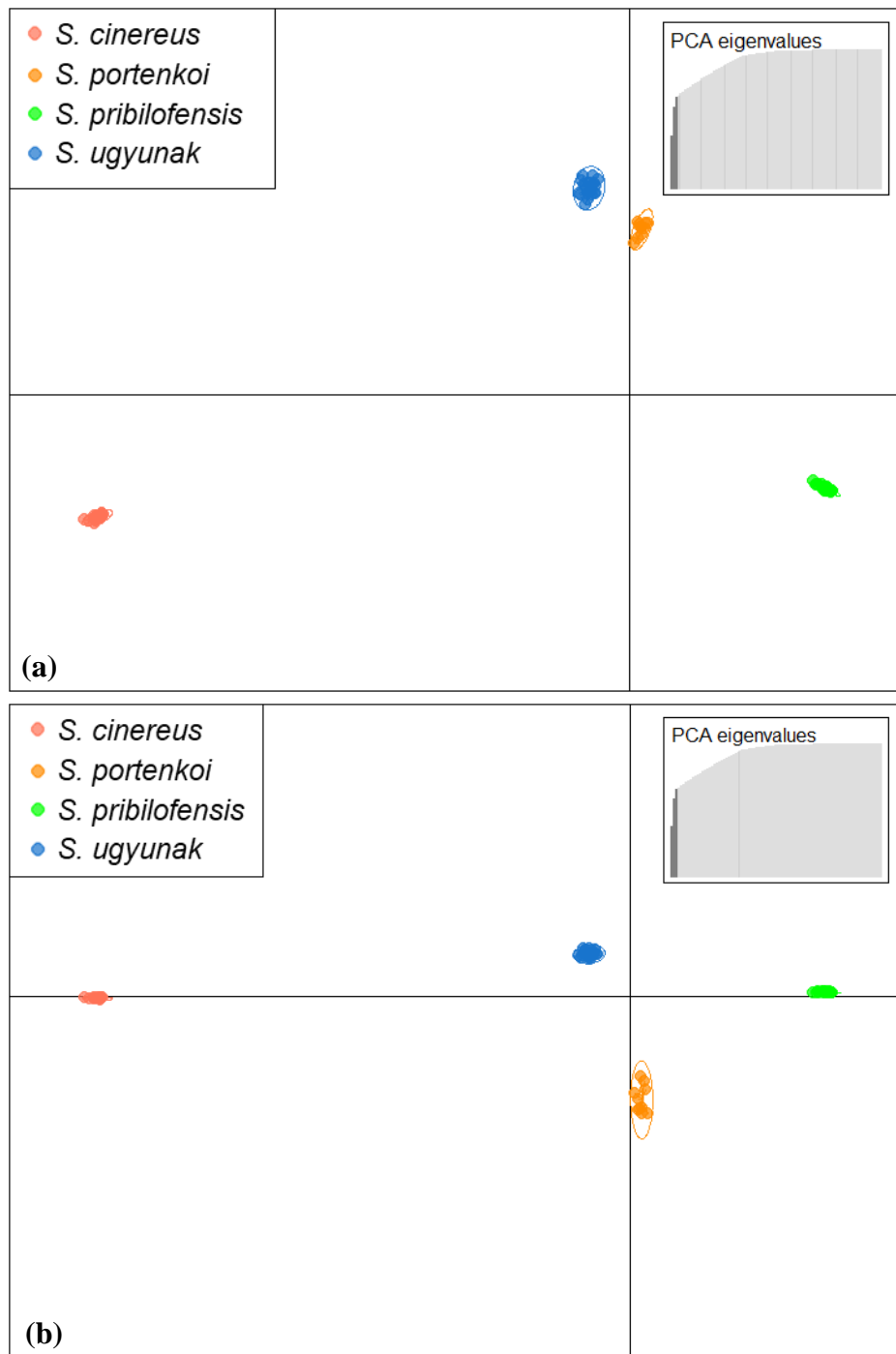


Figure 3.3 Neutral SNP DAPC

DAPC scatter plots based on the neutral SNP dataset. The cumulative variance explained by the 3 principle components retained for this analysis is shown in the top right of each plot. **(a)** DAPC using the first and second principle components; **(b)** DAPC using the first and third principle components.

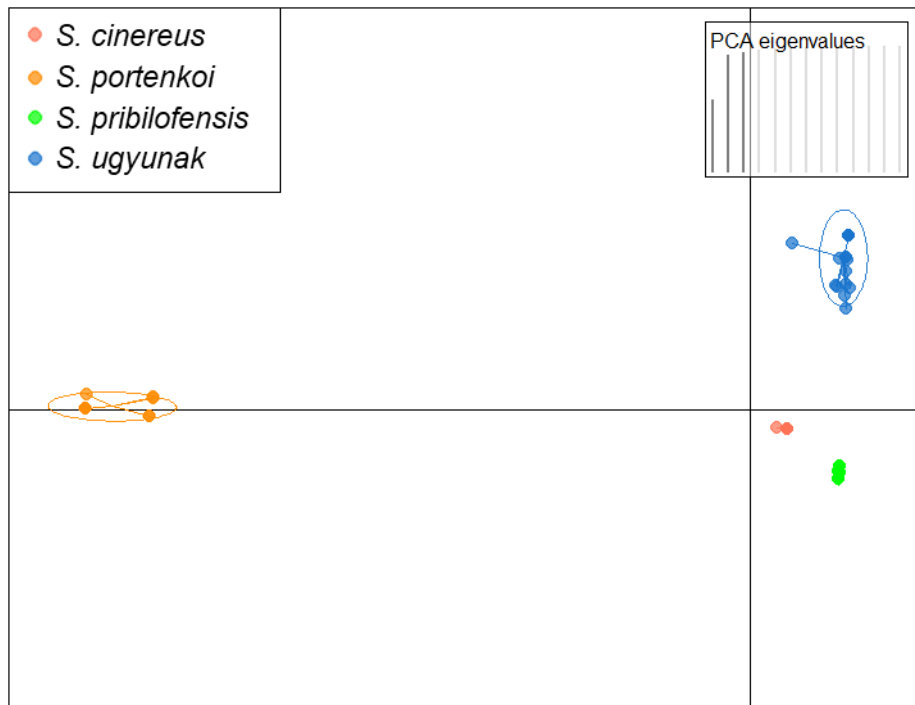


Figure 3.4 Non-neutral SNP DAPC

DAPC scatter plot based on the non-neutral SNP dataset. The cumulative variance explained by the 3 principle components retained for this analysis is shown in the top right of the plot. The individuals are plotted using the first and second principle components.

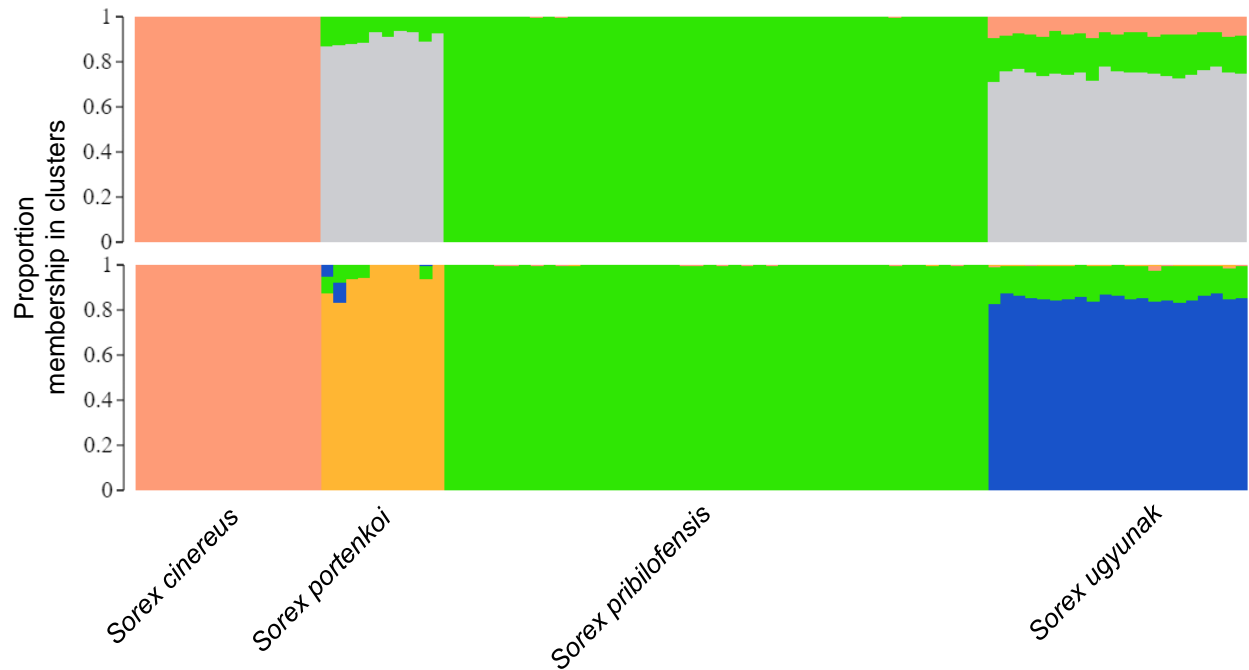


Figure 3.5 Neutral SNP Structure plots

Structure plots based on the neutral SNP dataset. Each color corresponds to a distinct genetic cluster, and each bar represents the proportion of each individual's genotype to each cluster. **(top)** Structure plot for most likely number of genetic clusters, K=3; **(bottom)** Structure plot for second most likely number of genetic clusters, K=4.

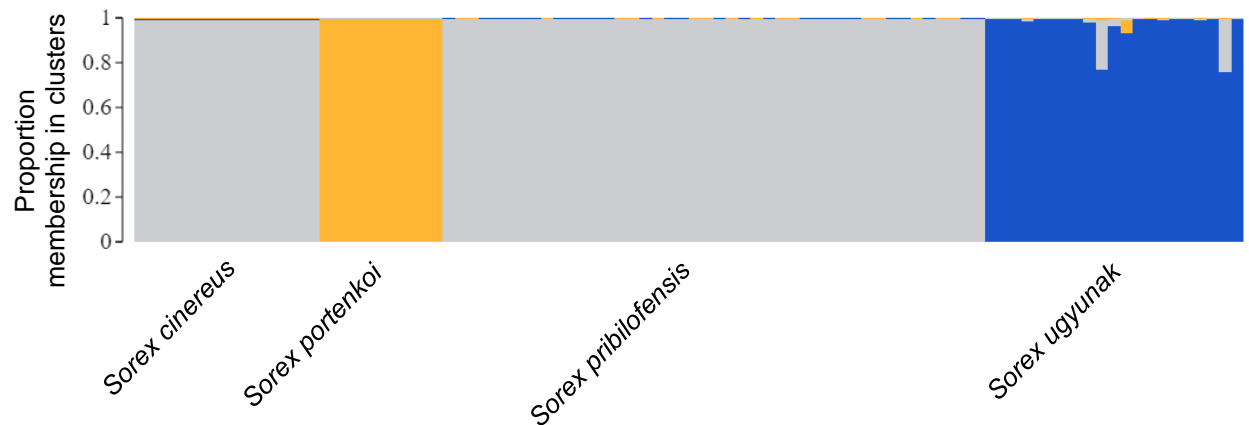


Figure 3.6 Non-neutral SNP Structure plot

Structure plot based on the non-neutral SNP dataset for the most likely number of genetic clusters, K=3. Each color corresponds to a distinct genetic cluster, and each bar represents the proportion of each individual's genotype to each cluster.

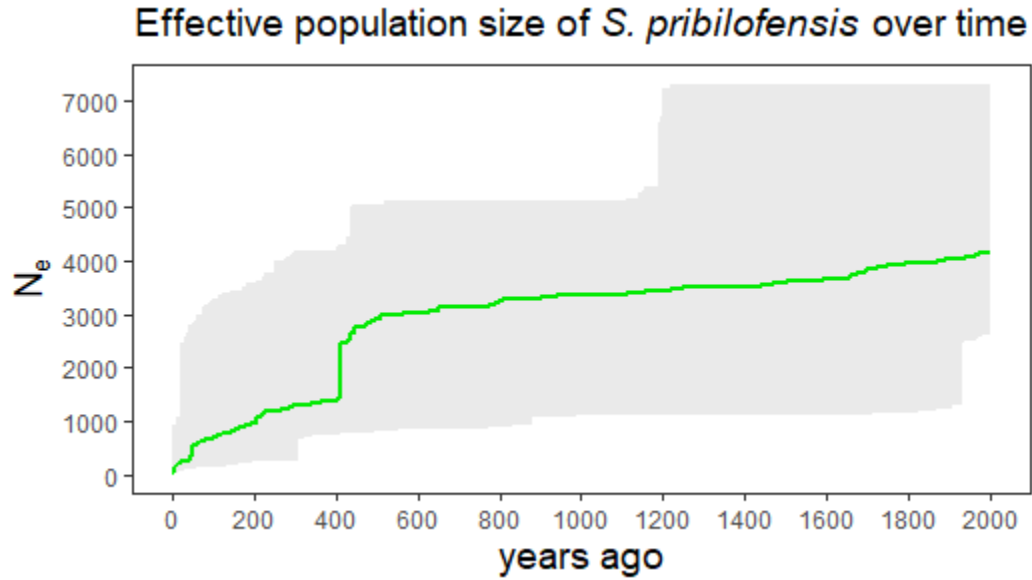


Figure 3.7 *S. pribilofensis* stairway plot

The effective population size of *S. pribilofensis* since 2,000 years ago, inferred with Stairway Plot v2.0. The grey shaded region indicates the 95% confidence interval.

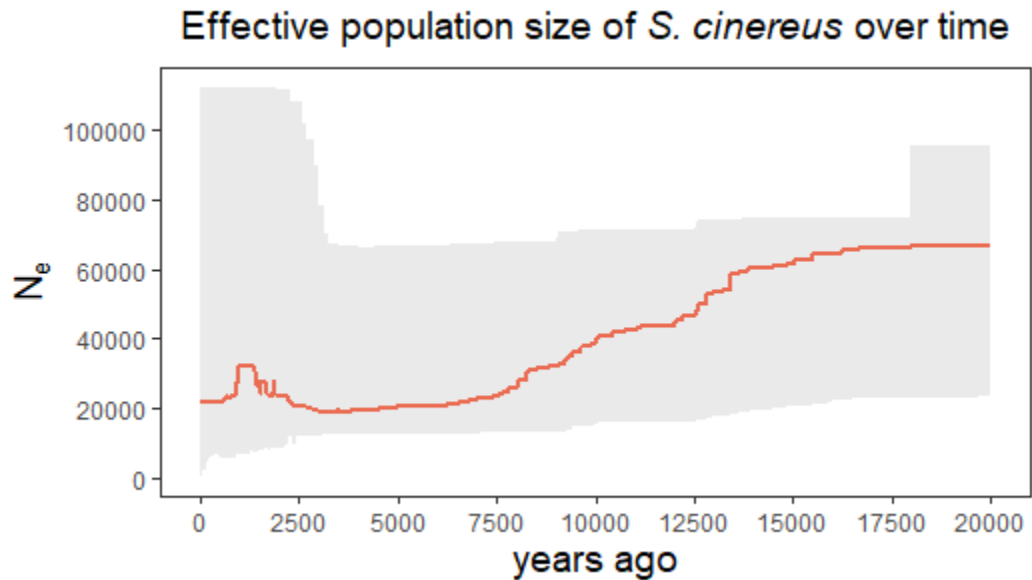


Figure 3.8 *S. cinereus* stairway plot

The effective population size of *S. cinereus* since 20,000 years ago, inferred with Stairway Plot v2.0. The grey shaded region indicates the 95% confidence interval.

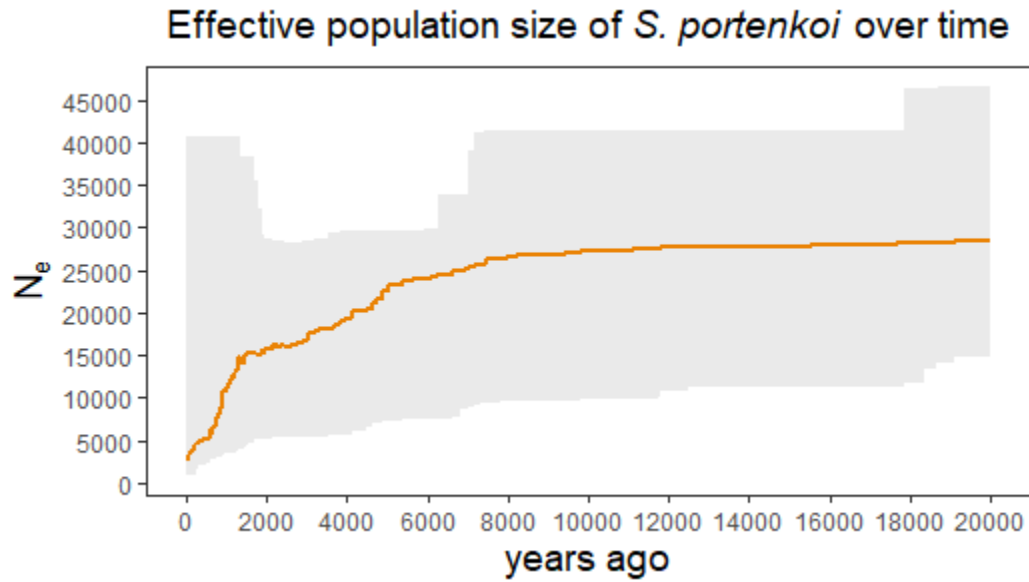


Figure 3.9 *S. portenkoi* stairway plot

The effective population size of *S. portenkoi* since 20,000 years ago, inferred with Stairway Plot v2.0. The grey shaded region indicates the 95% confidence interval.

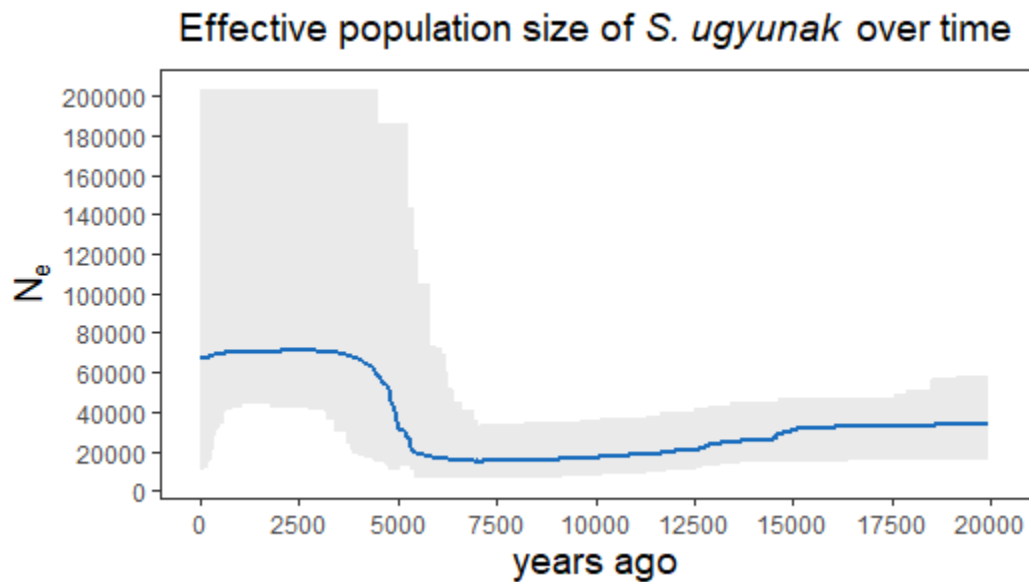


Figure 3.10 *S. ugyunak* stairway plot

The effective population size of *S. ugyunak* since 20,000 years ago, inferred with Stairway Plot v2.0. The grey shaded region indicates the 95% confidence interval.

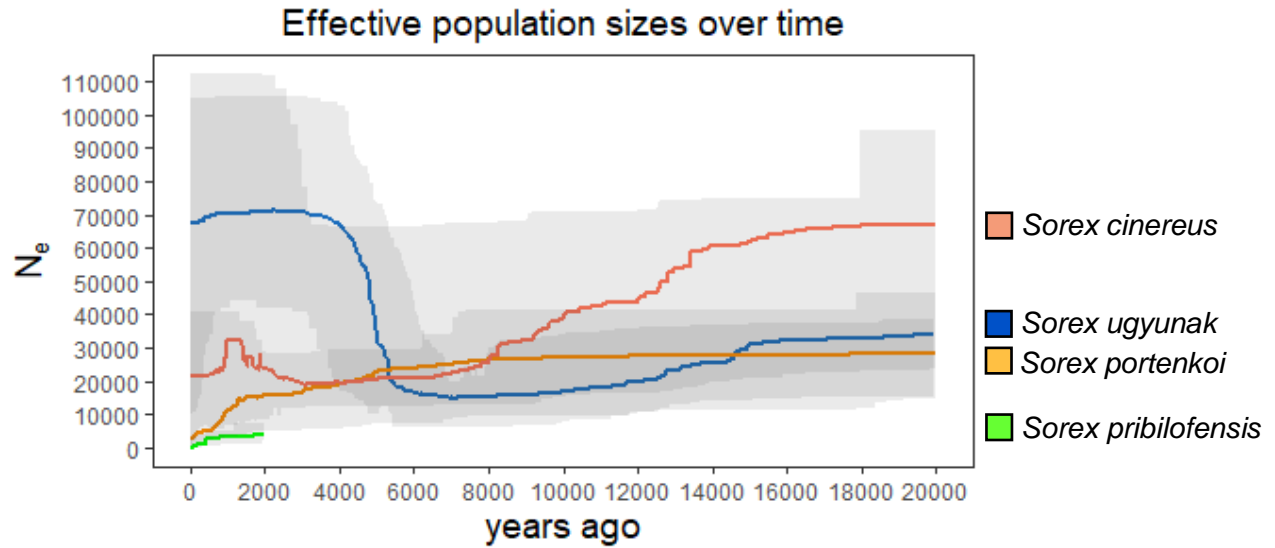


Figure 3.11 Combined stairway plots

The effective population sizes of the four species represented by the SNP dataset since 20,000 years ago, inferred with Stairway Plot v2.0. The grey shaded regions indicate the 95% confidence interval for each species. To facilitate comparison on the same plot, the upper bound for *S. ugyunak* is the 75% confidence interval.

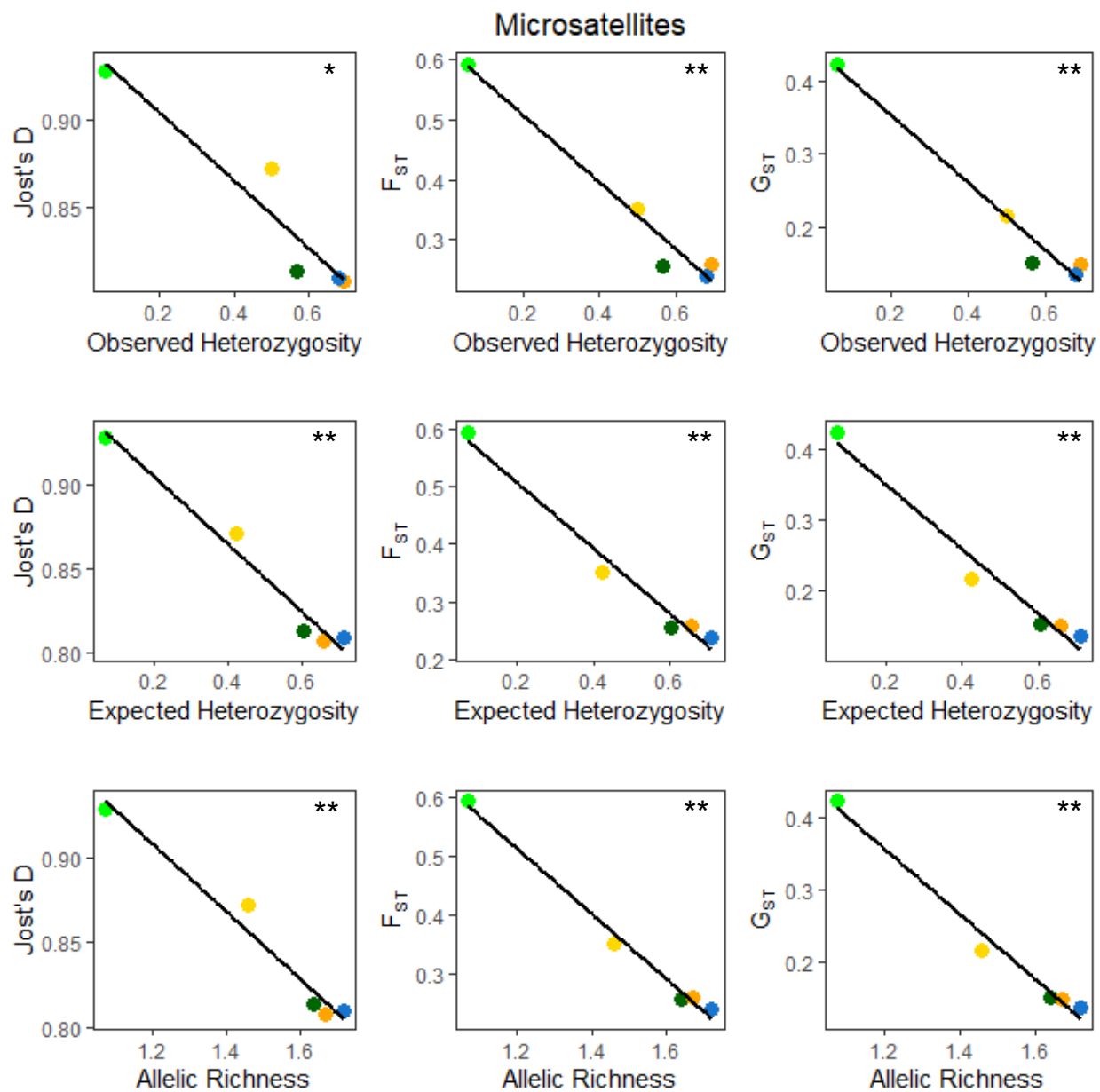
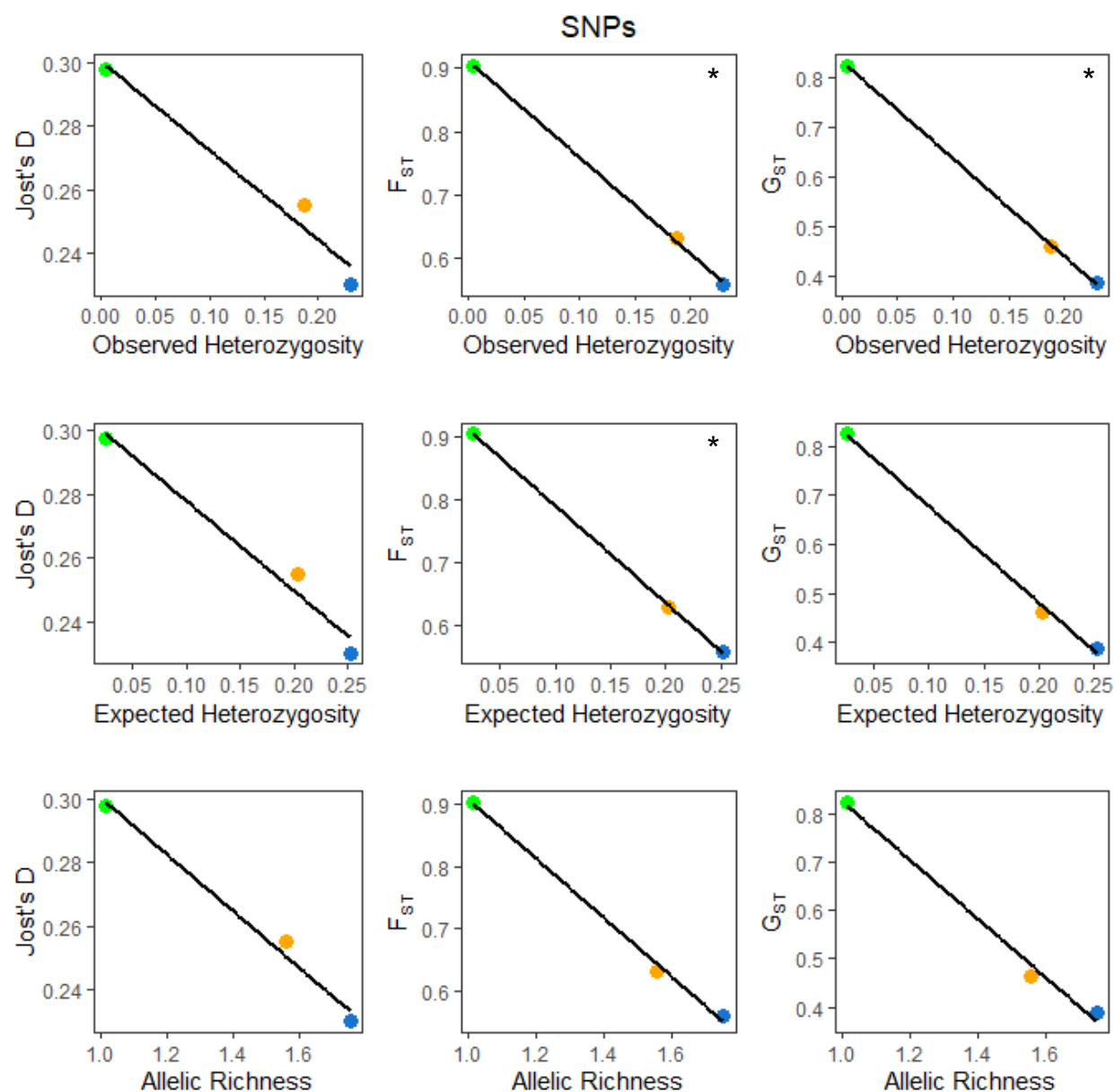


Figure 3.12 Microsatellite linear regressions

Plots of linear regressions between multiple measures of genetic variability and multiple measures of genetic differentiation for the microsatellite dataset.

*= $p < 0.05$, **= $p < 0.01$



Species: ● *S. portenkoi* ● *S. pribilofensis* ● *S. ugyunak*

Figure 3.13 Neutral SNP linear regressions

Plots of linear regressions between multiple measures of genetic variability and multiple measures of genetic differentiation for the neutral SNP dataset.

*= $p < 0.05$

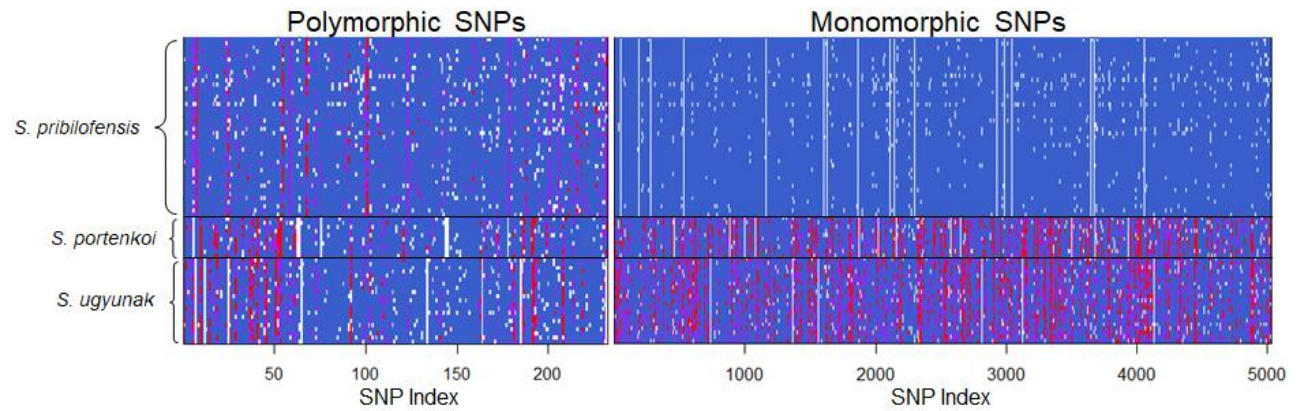


Figure 3.14 Polymorphic and monomorphic SNPs within *S. pribilofensis*

Allele frequencies within the neutral SNP dataset at the 232 polymorphic loci within *S. pribilofensis* and the 5,031 monomorphic loci within *S. pribilofensis*. A random subset of SNPs for each kind of locus is plotted along the x-axis, and individual shrews, grouped by species, are plotted along the y-axis. Colors represent the genotype of each shrew at each SNP; blue= homozygosity for the first allele (AA), purple= heterozygosity (Aa), red= homozygosity for the second allele (aa).

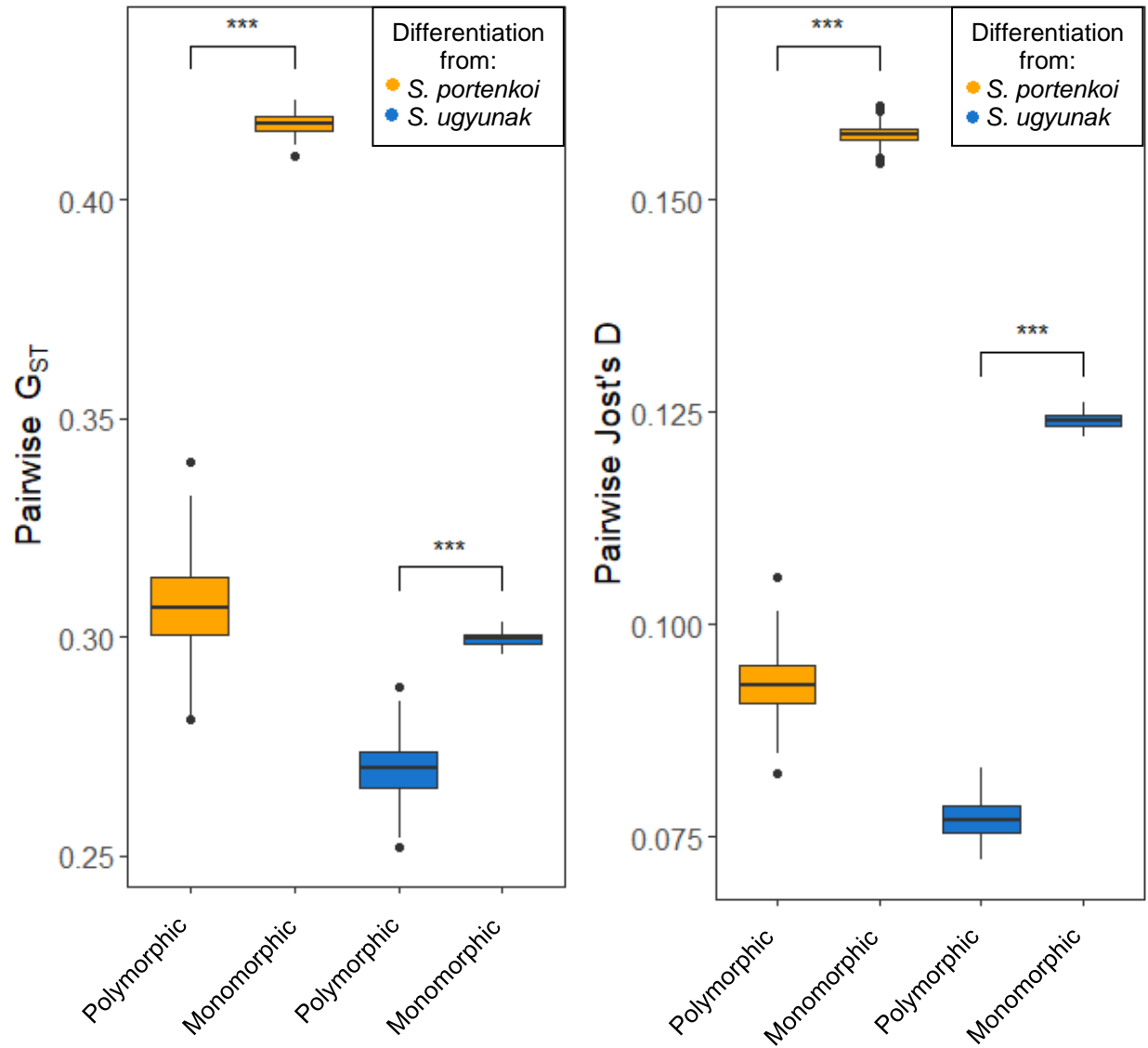


Figure 3.15 Pairwise genetic differentiation boxplots

Pairwise genetic differentiation between *S. pribilofensis* and *S. portenkoi* (orange), and *S. pribilofensis* and *S. ugyunak* (blue), with loci split by whether they are polymorphic or monomorphic within *S. pribilofensis*. The median of each boxplot is concurrent with the point estimate obtained for the neutral SNP dataset and the variance was estimated through 200 bootstrap samples. *** $p < 0.001$

Tables

Table 3.1 Species of the *Sorex cinereus* complex

Major clades and distribution of shrews of the *Sorex cinereus* species complex.

Species	Biogeographic Realm	Distribution
Beringian clade		
<i>S. camtschatica</i>	Palearctic	Magadanskaya and Kamchatka Peninsula (Russia)
<i>S. leucogaster</i>	Palearctic	Paramushir Island
<i>S. portenkoi</i>	Palearctic	Chukotka (Russia)
<i>S. pribilofensis</i>	Bering Strait	St. Paul Island
<i>S. jacksoni</i>	Bering Strait	St. Lawrence Island
<i>S. ugyunak</i>	Nearctic	Northernmost Alaska and Canada
<i>S. haydeni</i>	Nearctic	Great Plains of Canada and USA
<i>S. preblei</i>	Nearctic	Columbia Basin/Plateau (USA)
<i>S. lyelli</i>	Nearctic	East Sierra Nevada mountains (California)
Southern clade		
<i>S. cinereus</i>	Nearctic	Widespread across Canada and USA
<i>S. emarginatus</i>	Nearctic	Sierra Madre Occidental (Mexico)
<i>S. milleri</i>	Nearctic	Sierra Madre Oriental (Mexico)
<i>S. longirostris</i>	Nearctic	Southeast USA

Table 3.2 Microsatellite effective population sizes

Effective population size (N_e) estimates for each species based on the microsatellite dataset.
 N=sample size, Harmonic mean N=sample size adjusted for missing genotypes, 95% CI= 95% confidence interval.

Species	N	Harmonic mean N	N_e	95% CI
<i>Sorex camtschatica</i>	5	4.3	∞	(2.9, ∞)
<i>Sorex cinereus</i>	14	14.0	∞	(∞ , ∞)
<i>Sorex jacksoni</i>	6	3.4	∞	(2.4, ∞)
<i>Sorex pribilofensis</i>	22	21.3	18.4	(1.9, ∞)
<i>Sorex portenkoi</i>	13	12.9	69.8	(22.6, ∞)
<i>Sorex ugyunak</i>	19	19.0	∞	(983.3, ∞)

Table 3.3 Neutral SNP effective population sizes

Effective population size (N_e) estimates for each species based on the neutral SNP dataset.
 N=sample size, Harmonic mean N=sample size adjusted for missing genotypes, 95% CI= 95% confidence interval.

Species	N	Harmonic mean N	N_e	95% CI
<i>Sorex cinereus</i>	15	13.7	∞	(236.5, ∞)
<i>Sorex pribilofensis</i>	44	40.0	2106.5	(166.0, ∞)
<i>Sorex portenkoi</i>	10	9.2	21.9	(9.0, 131.4)
<i>Sorex ugyunak</i>	21	19.5	∞	(∞ , ∞)

Table 3.4 Genetic variability: microsatellites and cytochrome b

Measures of genetic variability for the microsatellite and cytochrome b datasets. H_o =observed heterozygosity, H_e =expected heterozygosity, Ar =allelic richness, F_{IS} =inbreeding coefficient, π =nucleotide diversity, H_d =haplotype diversity.

Species	microsatellites				cytochrome b	
	H_o	H_e	Ar	F_{IS}	π	H_d
<i>S. camtschatica</i>	0.565	0.604	1.639	0.186	0.0023	0.714
<i>S. cinereus</i>	0.634	0.681	1.694	0.106	0.0031	0.967
<i>S. jacksoni</i>	0.500	0.424	1.459	-0.051	0.0019	0.524
<i>S. portenkoi</i>	0.691	0.659	1.672	-0.008	0.0004	0.385
<i>S. pribilofensis</i>	0.059	0.071	1.072	0.195	0.0001	0.133
<i>S. ugyunak</i>	0.678	0.711	1.720	0.073	0.0017	0.789

Table 3.5 Genetic variability: SNPs

Measures of genetic variability for the total SNP dataset. H_o =observed heterozygosity, H_e =expected heterozygosity, Ar =allelic richness, π =nucleotide diversity, F_{IS} =inbreeding coefficient.

Species	SNPs				
	H_o	H_e	Ar	π	F_{IS}
<i>S. camtschatica</i>	-	-	-	-	-
<i>S. cinereus</i>	0.1453	0.315	1.496	0.1749	0.0877
<i>S. jacksoni</i>	-	-	-	-	-
<i>S. portenkoi</i>	0.0887	0.112	1.264	0.0927	0.0122
<i>S. pribilofensis</i>	0.0029	0.033	1.008	0.0021	-0.0014
<i>S. ugyunak</i>	0.1071	0.137	1.355	0.1152	0.0293

Table 3.6 Genetic variability: neutral SNPs

Measures of genetic variability for the 5,263 SNPs included in the neutral dataset. Measures for *S. cinereus* were not calculated as the neutral SNPs were identified using only the species structure between *S. pribilofensis*, *S. portenkoi*, and *S. ugyunak*. H_o =observed heterozygosity, H_e =expected heterozygosity, A_r =allelic richness, π =nucleotide diversity.

Species	Neutral SNPs			
	H_o	H_e	A_r	π
<i>S. portenkoi</i>	0.1873	0.203	1.559	0.1985
<i>S. pribilofensis</i>	0.0048	0.026	1.016	0.0040
<i>S. ugyunak</i>	0.2290	0.252	1.752	0.2484

Table 3.7 Genetic variability: polymorphic vs monomorphic SNPs

Measures of genetic variability at the 232 polymorphic SNPs within *S. pribilofensis* and the 5,031 monomorphic SNPs within *S. pribilofensis*. H_o =observed heterozygosity, H_e =expected heterozygosity, A_r =allelic richness, π =nucleotide diversity.

Species	Polymorphic within <i>S. pribilofensis</i>				Monomorphic within <i>S. pribilofensis</i>			
	H_o	H_e	A_r	π	H_o	H_e	A_r	π
<i>Sorex portenkoi</i>	0.098	0.122	1.291	0.101	0.191	0.207	1.571	0.203
<i>Sorex pribilofensis</i>	0.106	0.087	1.365	0.088	0	0.023	1	0
<i>Sorex ugyunak</i>	0.105	0.134	1.333	0.105	0.235	0.257	1.771	0.255

References

- Adamack, A. T., & Gruber, B. (2014). POP GEN REPORT: Simplifying basic population genetic analyses in R. *Methods in Ecology and Evolution*, 5(4), 384–387.
<https://doi.org/10.1111/2041-210X.12158>
- Agashe, D., Falk, J. J., & Bolnick, D. I. (2011). Effects of founding genetic variation on adaptation to a novel resource: genetic diversity facilitates adaptive niche shift. *Evolution*, 65(9), 2481–2491. <https://doi.org/10.1111/j.1558-5646.2011.01307.x>
- Avise, J. C., Walker, D., & Johns, G. C. (1998). Speciation durations and Pleistocene effects on vertebrate phylogeography. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1407), 1707–1712. <https://doi.org/10.1098/rspb.1998.0492>
- Barnosky, A. D. (2005). Effects of Quaternary Climatic Change on Speciation in Mammals. *Journal of Mammalian Evolution*, 12(1–2), 247–264. <https://doi.org/10.1007/s10914-005-4858-8>
- Baum, D. A., & Donoghue, M. J. (1995). Choosing among Alternative “Phylogenetic” Species Concepts. *Systematic Botany*, 20(4), 560. <https://doi.org/10.2307/2419810>
- Benazzo, A., Trucchi, E., Cahill, J. A., Maisano Delser, P., Mona, S., Fumagalli, M., Bunnefeld, L., Cornetti, L., Ghirotto, S., Girardi, M., Ometto, L., Panziera, A., Rota-Stabelli, O., Zanetti, E., Karamanlidis, A., Groff, C., Paule, L., Gentile, L., Vilà, C., ... Bertorelle, G. (2017). Survival and divergence in a small group: The extraordinary genomic history of the endangered Apennine brown bear stragglers. *Proceedings of the National Academy of Sciences*, 114(45), E9589–E9597. <https://doi.org/10.1073/pnas.1707279114>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B*, 57, 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Beynon, S. E., Slavov, G. T., Farré, M., Sunduimijid, B., Waddams, K., Davies, B., Haresign, W., Kijas, J., MacLeod, I. M., Newbold, C. J., Davies, L., & Larkin, D. M. (2015). Population structure and history of the Welsh sheep breeds determined by whole genome genotyping. *BMC Genetics*, 16(1), 65. <https://doi.org/10.1186/s12863-015-0216-x>
- Bi, K., Linderroth, T., Singhal, S., Vanderpool, D., Patton, J. L., Nielsen, R., Moritz, C., & Good, J. M. (2019). Temporal genomic contrasts reveal rapid evolutionary responses in an alpine mammal during recent climate change. *PLOS Genetics*, 15(5), e1008119. <https://doi.org/10.1371/journal.pgen.1008119>
- Bintanja, R., van de Wal, R. S. W., & Oerlemans, J. (2005). Modelled atmospheric temperatures and global sea levels over the past million years. *Nature*, 437(7055), 125–128. <https://doi.org/10.1038/nature03975>
- Bolger, A. M., Lohse, M., Usadel, B., Planck, M., Plant, M., & Mühlenberg, A. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 1–7. <https://doi.org/10.1093/bioinformatics/btu170>
- Bouckaert, R. R., & Drummond, A. J. (2017). bModelTest: Bayesian phylogenetic site model averaging and model comparison. *BMC Evolutionary Biology*, 17(1), 42. <https://doi.org/10.1186/s12862-017-0890-6>
- Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., De Maio, N., Matschiner, M., Mendes, F. K., Müller, N. F., Ogilvie, H. A., du Plessis, L., Poppinga, A., Rambaut, A., Rasmussen, D., Siveroni, I., ... Drummond, A. J. (2019). BEAST 2.5: An advanced software platform for Bayesian

- evolutionary analysis. *PLOS Computational Biology*, 15(4), e1006650.
<https://doi.org/10.1371/journal.pcbi.1006650>
- Brigham-Grette, J., Gualtieri, L. M., Glushkova, O. Y., Hamilton, T. D., Mostoller, D., & Kotov, A. (2003). Chlorine-36 and ^{14}C chronology support a limited last glacial maximum across central Chukotka, northeastern Siberia, and no Beringian ice sheet. *Quaternary Research*, 59(3), 386–398. [https://doi.org/10.1016/S0033-5894\(03\)00058-9](https://doi.org/10.1016/S0033-5894(03)00058-9)
- Brown, J. H. (1978). The theory of insular biogeography and the distribution of boreal birds and mammals. *Great Basin Naturalist Memoirs*, 2, 20.
- Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A., & RoyChoudhury, A. (2012). Inferring Species Trees Directly from Biallelic Genetic Markers: Bypassing Gene Trees in a Full Coalescent Analysis. *Molecular Biology and Evolution*, 29(8), 1917–1932.
<https://doi.org/10.1093/molbev/mss086>
- Byrd, G. V., & Norvell, N. (1993). Status of the Pribilof Shrew Based on Summer Distribution and Habitat Use. *Northwestern Naturalist*, 74(2), 49. <https://doi.org/10.2307/3536793>
- Callaghan, T. V., Björn, L. O., Chernov, Y., Chapin, T., Christensen, T. R., Huntley, B., Ims, R. A., Johansson, M., Jolly, D., Jonasson, S., Matveyeva, N., Panikov, N., Oechel, W., & Shaver, G. (2004). Past Changes in Arctic Terrestrial Ecosystems, Climate and UV Radiation. *AMBIO: A Journal of the Human Environment*, 33(7), 398–403.
<https://doi.org/10.1579/0044-7447-33.7.398>
- Campbell, C. D., Chong, J. X., Malig, M., Ko, A., Dumont, B. L., Han, L., Vives, L., O’Roak, B. J., Sudmant, P. H., Shendure, J., Abney, M., Ober, C., & Eichler, E. E. (2012). Estimating the human mutation rate using autozygosity in a founder population. *Nature Genetics*, 44(11), 1277–1281. <https://doi.org/10.1038/ng.2418>
- Ceballos, G., Ehrlich, P. R., Barnosky, A. D., García, A., Pringle, R. M., & Palmer, T. M. (2015). Accelerated modern human-induced species losses: Entering the sixth mass extinction. *Science Advances*, 1(5), e1400253. <https://doi.org/10.1126/sciadv.1400253>
- Chakraborty, R., & Nei, M. (1977). Bottleneck Effects on Average Heterozygosity and Genetic Distance with the Stepwise Mutation Model. *Evolution*, 31(2), 347–356.
- Charlesworth, B. (2009). Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, 10(3), 195–205. <https://doi.org/10.1038/nrg2526>
- Chifman, J., & Kubatko, L. (2014). Quartet Inference from SNP Data Under the Coalescent Model. *Bioinformatics*, 30(23), 3317–3324.
<https://doi.org/10.1093/bioinformatics/btu530>
- Clegg, S. M., Degnan, S. M., Moritz, C., Estoup, A., Kikkawa, J., & Owens, I. P. F. (2002). Microevolution in island forms: the roles of drift and directional selection in morphological divergence of a passerine bird. *Evolution*, 56(10), 2090–2099.
<https://doi.org/10.1111/j.0014-3820.2002.tb00134.x>
- Colella, J. P., Talbot, S. L., Brochmann, C., Taylor, E. B., Hoberg, E. P., & Cook, J. A. (2020). Conservation Genomics in a Changing Arctic. *Trends in Ecology & Evolution*, 35(2), 149–162. <https://doi.org/10.1016/j.tree.2019.09.008>
- Colinvaux, P. (1981). Historical Ecology in Beringia: The South Land Bridge Coast at St. Paul Island. *Quaternary Research*, 16(1), 18–36. [https://doi.org/10.1016/0033-5894\(81\)90125-3](https://doi.org/10.1016/0033-5894(81)90125-3)
- Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichet, T., Gao, X., Jr, W. J. G., Johns, T., Krinner, G., Shongwe, M., Weaver, A. J., Wehner, M., Allen, M. R., Andrews, T., Beyerle, U., Bitz, C. M., Bony, S., Booth, B. B. B., Brooks, H. E., ... Tett, S. (2013).

- Long-term Climate Change: Projections, Commitments and Irreversibility. *Climate Change 2013 - The Physical Science Basis: Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, 1029–1136.
- Combe, F. J., Sikes, D. S., Tkach, V. V., & Hope, A. G. (2021). Origins and diversity of the Bering Sea Island fauna: Shifting linkages across the northern continents. *Biodiversity and Conservation*. <https://doi.org/10.1007/s10531-021-02153-3>
- Comes, H. P., & Kadereit, J. W. (1998). The effect of Quaternary climatic changes on plant distribution and evolution. *Trends in Plant Science*, 3(11), 432–438. [https://doi.org/10.1016/S1360-1385\(98\)01327-2](https://doi.org/10.1016/S1360-1385(98)01327-2)
- Cook, J. A., Galbreath, K. E., Bell, K. C., Campbell, M. L., Carrière, S., Colella, J. P., Dawson, N. G., Dunnum, J. L., Eckerlin, R. P., Fedorov, V., Greiman, S. E., Haas, G. M. S., Haukialmi, V., Henttonen, H., Hope, A. G., Jackson, D., Jung, T. S., Koehler, A. V., Kinsella, J. M., ... Hoberg, E. P. (2016). The Beringian Coevolution Project: Holistic collections of mammals and associated parasites reveal novel perspectives on evolutionary and environmental change in the North. *Arctic Science*, 3(3), 585–617. <https://doi.org/10.1139/as-2016-0042>
- Coyne, J. A., & Orr, H. A. (1989). Patterns of speciation in drosophila. *Evolution*, 43(2), 362–381. <https://doi.org/10.1111/j.1558-5646.1989.tb04233.x>
- Dale Guthrie, R. (2004). Radiocarbon evidence of mid-Holocene mammoths stranded on an Alaskan Bering Sea island. *Nature*, 429(6993), 746–749. <https://doi.org/10.1038/nature02612>
- de Queiroz, K. (2005). A Unified Concept of Species and Its Consequences for the Future of Taxonomy. *Proceedings of the California Academy of Sciences*.
- Demboski, J. R., & Cook, J. A. (2003). Phylogenetic diversification within the *Sorex cinereus* group (soricidae). *Journal of Mammalogy*, 84(1), 15.
- Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., & Ovenden, J. R. (2014). NEESTIMATOR v2: Re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Molecular Ecology Resources*, 14(1), 209–214. <https://doi.org/10.1111/1755-0998.12157>
- Dobzhansky, T., & Pavlovsky, O. (1957). An Experimental Study of Interaction between Genetic Drift and Natural Selection. *Evolution*, 11(3), 311. <https://doi.org/10.2307/2405795>
- Dokuchaev, N. E. (1989). Population ecology of *Sorex* shrews in North-East Siberia. *Annales Zoologici Fennici*, 26(4), 371–379.
- Earl, D. A., & vonHoldt, B. M. (2012). Structure Harvester: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4(2), 359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Elias, S. A., & Brigham-Grette, J. (2013). Glaciations | Late Pleistocene Glacial Events in Beringia. In *Encyclopedia of Quaternary Science* (pp. 191–201). Elsevier. <https://doi.org/10.1016/B978-0-444-53643-3.00116-3>
- Ellegren, H. (2014). Genome sequencing and population genomics in non-model organisms. *Trends in Ecology & Evolution*, 29(1), 51–63. <https://doi.org/10.1016/j.tree.2013.09.008>
- Ely, C. V., Bordignon, S. A. de L., Trevisan, R., & Boldrini, I. I. (2017). Implications of poor taxonomy in conservation. *Journal for Nature Conservation*, 36, 10–13. <https://doi.org/10.1016/j.jnc.2017.01.003>

- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. *Molecular Ecology*, 14(8), 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Excoffier, L., Foll, M., & Petit, R. J. (2009). Genetic Consequences of Range Expansions. *Annual Review of Ecology, Evolution, and Systematics*, 40(1), 481–501. <https://doi.org/10.1146/annurev.ecolsys.39.110707.173414>
- Fay, F. H., & Sease, J. I. (1985). *Preliminary Status Survey of Selected Small Mammals*. 53.
- Fišer, C., Robinson, C. T., & Malard, F. (2018). Cryptic species as a window into the paradigm shift of the species concept. *Molecular Ecology*, 27(3), 613–635. <https://doi.org/10.1111/mec.14486>
- Fisher, Ronald Aylmer. (1958). *The genetical theory of natural selection*. Clarendon Press, Oxford, UK.
- Foll, M., & Gaggiotti, O. (2008). A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant and Codominant Markers: A Bayesian Perspective. *Genetics*, 180(2), 977–993. <https://doi.org/10.1534/genetics.108.092221>
- Frankham, R. (1997b). Do island populations have less genetic variation than mainland populations? *Heredity*, 78(3), 311–327. <https://doi.org/10.1038/hdy.1997.46>
- Funk, W. C., Lovich, R. E., Hohenlohe, P. A., Hofman, C. A., Morrison, S. A., Sillett, T. S., Ghalambor, C. K., Maldonado, J. E., Rick, T. C., Day, M. D., Polato, N. R., Fitzpatrick, S. W., Coonan, T. J., Crooks, K. R., Dillon, A., Garcelon, D. K., King, J. L., Boser, C. L., Gould, N., & Andelt, W. F. (2016). Adaptive divergence despite strong genetic drift: Genomic analysis of the evolutionary mechanisms causing genetic differentiation in the island fox (*Urocyon littoralis*). *Molecular Ecology*, 25(10), 2176–2194. <https://doi.org/10.1111/mec.13605>
- Galbreath, K. E., Hoberg, E. P., Cook, J. A., Armién, B., Bell, K. C., Campbell, M. L., Dunnum, J. L., Dursahinhan, A. T., Eckerlin, R. P., Gardner, S. L., Greiman, S. E., Henttonen, H., Jiménez, F. A., Koehler, A. V. A., Nyamsuren, B., Tkach, V. V., Torres-Pérez, F., Tsvetkova, A., & Hope, A. G. (2019). Building an integrated infrastructure for exploring biodiversity: Field collections and archives of mammals and parasites. *Journal of Mammalogy*, 100(2), 382–393. <https://doi.org/10.1093/jmammal/gyz048>
- Giglio, R. M., Rocke, T. E., Osorio, J. E., & Latch, E. K. (2020). Characterizing patterns of genomic variation in the threatened Utah prairie dog: Implications for conservation and management. *Evolutionary Applications*, eua.13179. <https://doi.org/10.1111/eva.13179>
- Goudet, J. (2005). Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, 5(1), 184–186. <https://doi.org/10.1111/j.1471-8286.2004.00828.x>
- Graham, N. R., Gruner, D. S., Lim, J. Y., & Gillespie, R. G. (2017). Island ecology and evolution: Challenges in the Anthropocene. *Environmental Conservation*, 44(4), 323–335. <https://doi.org/10.1017/S0376892917000315>
- Graham, R. W., Belmecheri, S., Choy, K., Culleton, B. J., Davies, L. J., Froese, D., Heintzman, P. D., Hritz, C., Kapp, J. D., Newsom, L. A., Rawcliffe, R., Saulnier-Talbot, É., Shapiro, B., Wang, Y., Williams, J. W., & Wooller, M. J. (2016). Timing and causes of mid-Holocene mammoth extinction on St. Paul Island, Alaska. *Proceedings of the National Academy of Sciences*, 113(33), 9310–9314. <https://doi.org/10.1073/pnas.1604903113>
- Grant, P. R. (2002). Unpredictable Evolution in a 30-Year Study of Darwin’s Finches. *Science*, 296(5568), 707–711. <https://doi.org/10.1126/science.1070315>

- Gruber, B., Unmack, P. J., Berry, O. F., & Georges, A. (2018). DARTR: An R package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Molecular Ecology Resources*, 18(3), 691–699. <https://doi.org/10.1111/1755-0998.12745>
- Guol, S. W., & Thompson, E. A. (1992). Performing the Exact Test of Hardy-Weinberg Proportion for Multiple Alleles. *Biometrics*, 48(2), 361–372. <https://doi.org/doi:10.2307/2532296>
- Hansen, C. C. R., Hvilsom, C., Schmidt, N. M., Aastrup, P., Van Coeverden de Groot, P. J., Siegismund, H. R., & Heller, R. (2018). The Muskox Lost a Substantial Part of Its Genetic Diversity on Its Long Road to Greenland. *Current Biology*, 28(24), 4022–4028.e5. <https://doi.org/10.1016/j.cub.2018.10.054>
- Heaney, L. R. (2000). Dynamic disequilibrium: A long-term, large-scale perspective on the equilibrium model of island biogeography: Long-term dynamics in island biogeography. *Global Ecology and Biogeography*, 9(1), 59–74. <https://doi.org/10.1046/j.1365-2699.2000.00163.x>
- Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, 405(6789), 907–913. <https://doi.org/10.1038/35016000>
- Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359(1442), 183–195. <https://doi.org/10.1098/rstb.2003.1388>
- Hoffmann, A. A., & Sgrò, C. M. (2011). Climate change and evolutionary adaptation. *Nature*, 470(7335), 479–485. <https://doi.org/10.1038/nature09670>
- Hohenlohe, P. A., Funk, W. C., & Rajora, O. P. (2021). Population genomics for wildlife conservation and management. *Molecular Ecology*, 30(1), 62–82. <https://doi.org/10.1111/mec.15720>
- Hope, A G, Greiman, S. E., Tkach, V. V., Hoberg, E. P., & Cook, J. A. (2016). Shrews and Their Parasites: Small Species Indicate Big Changes. *Arctic Report Card*, 81–91.
- Hope, A. G., Waltari, E., Malaney, J. L., Payer, D. C., Cook, J. A., & Talbot, S. L. (2015). Arctic biodiversity: Increasing richness accompanies shrinking refugia for a cold-associated tundra fauna. *Ecosphere*, 6(9), art159. <https://doi.org/10.1890/ES15-00104.1>
- Hope, Andrew G., & Elias, S. (2019). Arctic Tundra Mammals. In *Encyclopedia of the World's Biomes* (pp. 11770–11771). Elsevier. <https://doi.org/10.1016/B978-0-12-409548-9.11770-1>
- Hope, Andrew G, Sandercock, B. K., & Malaney, J. L. (2018). Collection of Scientific Specimens: Benefits for Biodiversity Sciences and Limited Impacts on Communities of Small Mammals. *BioScience*, 68(1), 35–42. <https://doi.org/10.1093/biosci/bix141>
- Hope, Andrew G., Speer, K. A., Demboski, J. R., Talbot, S. L., & Cook, J. A. (2012). A climate for speciation: Rapid spatial diversification within the *Sorex cinereus* complex of shrews. *Molecular Phylogenetics and Evolution*, 64(3), 671–684. <https://doi.org/10.1016/j.ympev.2012.05.021>
- Hope, Andrew G, Waltari, E., Dokuchaev, N. E., Abramov, S., Dupal, T., Tsvetkova, A., Henttonen, H., Macdonald, S. O., & Cook, J. A. (2010). High-latitude diversification within Eurasian least shrews and Alaska tiny shrews (Soricidae). *Journal of Mammalogy*, 91(5), 17.
- IUCN. 2016. The IUCN Red List of Threatened Species. Version 2018-1. Available at: www.iucnredlist.org. (Accessed: 3 May 2018).

- Jackson, H. H. T. (1928). A taxonomic review of the American long-tailed shrews: (genera *Sorex* and *Microsorex*). *North American Fauna*, 51, 1–238.
<https://doi.org/10.3996/nafa.51.0001>
- Jakobsson, M., Pearce, C., Cronin, T. M., Backman, J., Anderson, L. G., Barrientos, N., Björk, G., Coxall, H., de Boer, A., Mayer, L. A., Mörtz, C.-M., Nilsson, J., Rattray, J. E., Stranne, C., Semiletov, I., & O'Regan, M. (2017). Post-glacial flooding of the Bering Land Bridge dated to 11 cal ka BP based on new geophysical and sediment records. *Climate of the Past*, 13(8), 991–1005. <https://doi.org/10.5194/cp-13-991-2017>
- Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11), 1403–1405. <https://doi.org/10.1093/bioinformatics/btn129>
- Jones, M. R., & Good, J. M. (2016). Targeted capture in evolutionary and ecological genomics. *Molecular Ecology*, 25(1), 185–202. <https://doi.org/10.1111/mec.13304>
- Jong, M. J., Jong, J. F., Hoelzel, A. R., & Janke, A. (2021). SambaR: An R package for fast, easy and reproducible population-genetic analyses of biallelic SNP data sets. *Molecular Ecology Resources*, 21(4), 1369–1379. <https://doi.org/10.1111/1755-0998.13339>
- Jordan, M. A., & Snell, H. L. (2008). Historical fragmentation of islands and genetic drift in populations of Galápagos lava lizards (*Microlophus albemarlensis* complex). *Molecular Ecology*, 17(5), 1224–1237. <https://doi.org/10.1111/j.1365-294X.2007.03658.x>
- Kaiser-Bunbury, C. N., Traveset, A., & Hansen, D. M. (2010). Conservation and restoration of plant–animal mutualisms on oceanic islands. *Perspectives in Plant Ecology, Evolution and Systematics*, 12(2), 131–143. <https://doi.org/10.1016/j.ppees.2009.10.002>
- Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, 90(430), 773–795. <https://doi.org/10.1080/01621459.1995.10476572>
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., & Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28(12), 1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>
- Keller, L. (2002). Inbreeding effects in wild populations. *Trends in Ecology & Evolution*, 17(5), 230–241. [https://doi.org/10.1016/S0169-5347\(02\)02489-8](https://doi.org/10.1016/S0169-5347(02)02489-8)
- Knaus, B. J., & Grünwald, N. J. (2017). VCFR: A package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources*, 17(1), 44–53.
<https://doi.org/10.1111/1755-0998.12549>
- Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics*, 15(1), 356.
<https://doi.org/10.1186/s12859-014-0356-4>
- Kueffer, C., Drake, D. R., & Fernández-Palacios, J. M. (2014). Island biology: Looking towards the future. *Biology Letters*, 10(10), 20140719. <https://doi.org/10.1098/rsbl.2014.0719>
- Kumar, S., & Subramanian, S. (2002). Mutation rates in mammalian genomes. *Proceedings of the National Academy of Sciences*, 99(2), 803–808.
<https://doi.org/10.1073/pnas.022629899>
- Leaché, A. D., Banbury, B. L., Felsenstein, J., de Oca, A. nieta-Montes, & Stamatakis, A. (2015). Short Tree, Long Tree, Right Tree, Wrong Tree: New Acquisition Bias Corrections for Inferring SNP Phylogenies. *Systematic Biology*, 64(6), 1032–1047.
<https://doi.org/10.1093/sysbio/syv053>

- Leaché, A. D., & Bouckaert, R. R. (2018). *Species Trees and Species Delimitation with SNAPP: A Tutorial and Worked Example*. 17.
- Leaché, A. D., Fujita, M. K., Minin, V. N., & Bouckaert, R. R. (2014). Species Delimitation using Genome-Wide SNP Data. *Systematic Biology*, 63(4), 534–542. <https://doi.org/10.1093/sysbio/syu018>
- Leclerc, C., Courchamp, F., & Bellard, C. (2018). Insular threat associations within taxa worldwide. *Scientific Reports*, 8(1), 6393. <https://doi.org/10.1038/s41598-018-24733-0>
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14), 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
- Liu, S., Lorenzen, E. D., Fumagalli, M., Li, B., Harris, K., Xiong, Z., Zhou, L., Korneliussen, T. S., Somel, M., Babbitt, C., Wray, G., Li, J., He, W., Wang, Z., Fu, W., Xiang, X., Morgan, C. C., Doherty, A., O’Connell, M. J., ... Wang, J. (2014). Population Genomics Reveal Recent Speciation and Rapid Evolutionary Adaptation in Polar Bears. *Cell*, 157(4), 785–794. <https://doi.org/10.1016/j.cell.2014.03.054>
- Liu, X., & Fu, Y.-X. (2020). Stairway Plot 2: Demographic history inference with folded SNP frequency spectra. *Genome Biology*, 21(1), 280. <https://doi.org/10.1186/s13059-020-02196-9>
- Lomolino, Mark V. (2000). A call for a new paradigm of island biogeography: Island Paradigms. *Global Ecology and Biogeography*, 9(1), 1–6. <https://doi.org/10.1046/j.1365-2699.2000.00185.x>
- Lomolino, Mark V., Brown, J. H., & Davis, R. (1989). Island Biogeography of Montane Forest Mammals in the American Southwest. *Ecology*, 70(1), 180–194. <https://doi.org/10.2307/1938425>
- Losos, J. B. (1998). Contingency and Determinism in Replicated Adaptive Radiations of Island Lizards. *Science*, 279(5359), 2115–2118. <https://doi.org/10.1126/science.279.5359.2115>
- Luu, K., Bazin, E., & Blum, M. G. B. (2017). *pcadapt*: An R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, 17(1), 67–77. <https://doi.org/10.1111/1755-0998.12592>
- MacArthur, R. H., & Wilson, E. O. (1967). *The theory of island biogeography*. Princeton University Press, Princeton, NJ.
- MacLeod, I. M., Larkin, D. M., Lewin, H. A., Hayes, B. J., & Goddard, M. E. (2013). Inferring Demography from Runs of Homozygosity in Whole-Genome Sequence, with Correction for Sequence Errors. *Molecular Biology and Evolution*, 30(9), 2209–2223. <https://doi.org/10.1093/molbev/mst125>
- Marandel, F., Lorange, P., Berthelé, O., Trenkel, V. M., Waples, R. S., & Lamy, J.-B. (2019). Estimating effective population size of large marine populations, is it feasible? *Fish and Fisheries*, 20(1), 189–198.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, 17(1), 10. <https://doi.org/10.14806/ej.17.1.200>
- Matute, D. R., & Cooper, B. S. (2021). Comparative studies on speciation: 30 years since Coyne and Orr. *Evolution*, evo.14181. <https://doi.org/10.1111/evo.14181>
- McLean, B. S., Bell, K. C., Dunnum, J. L., Abrahamson, B., Colella, J. P., Deardorff, E. R., Weber, J. A., Jones, A. K., Salazar-Miralles, F., & Cook, J. A. (2016). Natural history collections-based research: Progress, promise, and best practices. *Journal of Mammalogy*, 97(1), 287–297. <https://doi.org/10.1093/jmammal/gyv178>

- Merriam, C. H. (1895). Synopsis of the American shrews of the genus *Sorex*. *North American Fauna*, 10, 57–125. <https://doi.org/10.3996/nafa.10.0003>
- Miller, A. D., Hoffmann, A. A., Tan, M. H., Young, M., Ahrens, C., Cocomazzo, M., Rattray, A., Ierodiaconou, D. A., Trembl, E., & Sherman, C. D. H. (2019). Local and regional scale habitat heterogeneity contribute to genetic adaptation in a commercially important marine mollusc (*Haliotis rubra*) from southeastern Australia. *Molecular Ecology*, 28(12), 3053–3072. <https://doi.org/10.1111/mec.15128>
- Miller, S. A., Dykes, D. D., & Polesky, H. F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research*, 16(3), 1215–1215. <https://doi.org/10.1093/nar/16.3.1215>
- Miskelly, C. M., & Powlesland, R. (2013). Conservation translocations of New Zealand birds, 1863–2012. *Notornis*, 60(1), 3–28.
- Moritz, C., Patton, J. L., Conroy, C. J., Parra, J. L., White, G. C., & Beissinger, S. R. (2008). Impact of a Century of Climate Change on Small-Mammal Communities in Yosemite National Park, USA. *Science*, 322(5899), 261–264. <https://doi.org/10.1126/science.1163428>
- Nei, M. (1987). *Molecular Evolutionary Genetics*. Columbia University Press.
- Nei, M., Maruyama, T., & Chakraborty, R. (1975). The Bottleneck Effect and Genetic Variability in Populations. *Evolution*, 1–10.
- Nichols, R. A., & Hewitt, G. M. (1994). The genetic consequences of long distance dispersal during colonization. *Heredity*, 72(3), 312–317. <https://doi.org/10.1038/hdy.1994.41>
- Nunziata, S. O., & Weisrock, D. W. (2018). Estimation of contemporary effective population size and population declines using RAD sequence data. *Heredity*, 120(3), 196–207.
- Ohta, T., & Tachida, H. (1990). Theoretical study of near neutrality. I. Heterozygosity and rate of mutant substitution. *Genetics*, 126(1), 219–229. <https://doi.org/10.1093/genetics/126.1.219>
- Ohta, Tomoko. (1992). The Nearly Neutral Theory of Molecular Evolution. *Annual Review of Ecology and Systematics*, 23(1), 263–286.
- Paradis, E. (2010). pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics*, 26(3), 419–420. <https://doi.org/10.1093/bioinformatics/btp696>
- Patterson, B. D. (1987). The Principle of Nested Subsets and Its Implications for Biological Conservation. *Conservation Biology*, 1(4), 323–334. <https://doi.org/10.1111/j.1523-1739.1987.tb00052.x>
- Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE*, 7(5), e37135. <https://doi.org/10.1371/journal.pone.0037135>
- Prentice, M. B., Bowman, J., Khidas, K., Koen, E. L., Row, J. R., Murray, D. L., & Wilson, P. J. (2017). Selection and drift influence genetic differentiation of insular Canada lynx (*Lynx canadensis*) on Newfoundland and Cape Breton Island. *Ecology and Evolution*, 7(9), 3281–3294. <https://doi.org/10.1002/ece3.2945>
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, 155(2), 945–959.
- Ramasamy, R., Ramasamy, S., Bindroo, B., & Naik, V. (2014). STRUCTURE PLOT: A program for drawing elegant STRUCTURE bar plots in user friendly interface. *SpringerPlus*, 3(1), 431. <https://doi.org/10.1186/2193-1801-3-431>

- Rambaut, A. (2018). *Tracer* (1.7.1) [Computer software]. <http://beast.community>
- Rambaut, A., & Drummond, A. J. (2012). *FigTree* (1.4.4) [Computer software].
- Rambaut, A., & Drummond, A. J. (2020). *TreeAnnotator* (2.6.2) [Computer software].
- Reed, D. H., & Frankham, R. (2003). Correlation between Fitness and Genetic Diversity. *Conservation Biology*, 17(1), 230–237. <https://doi.org/10.1046/j.1523-1739.2003.01236.x>
- Roach, J. C., Glusman, G., Smit, A. F. A., Huff, C. D., Hubley, R., Shannon, P. T., Rowen, L., Pant, K. P., Goodman, N., Bamshad, M., Shendure, J., Drmanac, R., Jorde, L. B., Hood, L., & Galas, D. J. (2010). Analysis of Genetic Inheritance in a Family Quartet by Whole-Genome Sequencing. *Science*, 328(5978), 636–639. <https://doi.org/10.1126/science.1186802>
- Robinson, J. A., Ortega-Del Vecchyo, D., Fan, Z., Kim, B. Y., vonHoldt, B. M., Marsden, C. D., Lohmueller, K. E., & Wayne, R. K. (2016). Genomic Flatlining in the Endangered Island Fox. *Current Biology*, 26(9), 1183–1189. <https://doi.org/10.1016/j.cub.2016.02.062>
- Rochette, N. C., & Catchen, J. M. (2017). Deriving genotypes from RAD-seq short-read data using Stacks. *Nature Protocols*, 12(12), 2640–2659. <https://doi.org/10.1038/nprot.2017.123>
- Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, 28(21), 4737–4754. <https://doi.org/10.1111/mec.15253>
- Roy, M. S., Geffen, E., Smith, D., Ostrander, E. A., & Wayne, R. K. (1994). Patterns of differentiation and hybridization in North American wolflike canids, revealed by analysis of microsatellite loci. *Molecular Biology and Evolution*, 11(4), 553–570. <https://doi.org/10.1093/oxfordjournals.molbev.a040137>
- Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sánchez-Gracia, A. (2017). DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. *Molecular Biology and Evolution*, 34(12), 3299–3302. <https://doi.org/10.1093/molbev/msx248>
- RStudio Team. (2020). *RStudio: Integrated Development Environment for R*. RStudio, PBC, Boston, MA. <http://www.rstudio.com/>
- Russell, J. C., & Kueffer, C. (2019). Island Biodiversity in the Anthropocene. *Annual Review of Environment and Resources*, 44(1), 31–60. <https://doi.org/10.1146/annurev-environ-101718-033245>
- Sher, A. (1999). Traffic lights at the Beringian crossroads. *Nature*, 397(6715), 103–104.
- Soltis, P. S., & Gitzendanner, M. A. (1999). Molecular Systematics and the Conservation of Rare Species. *Conservation Biology*, 13(3), 471–483. <https://doi.org/10.1046/j.1523-1739.1999.97286.x>
- Sonsthagen, S. A., Sage, G. K., Fowler, M., Hope, A. G., Cook, J. A., & Talbot, S. L. (2013). Development and characterization of 21 polymorphic microsatellite markers for the barren-ground shrew, *Sorex ugyunak* (Mammalia: Soricidae), through next-generation sequencing, and cross-species amplification in the masked shrew, *S. cinereus*. *Conservation Genetics Resources*, 5(2), 315–318. <https://doi.org/10.1007/s12686-012-9792-5>
- Spalding, M. D., Fox, H. E., Allen, G. R., Davidson, N., Ferdaña, Z. A., Finlayson, M., Halpern, B. S., Jorge, M. A., Lombana, A., Lourie, S. A., Martin, K. D., McManus, E., Molnar, J., Recchia, C. A., & Robertson, J. (2007). Marine Ecoregions of the World: A

- Bioregionalization of Coastal and Shelf Areas. *BioScience*, 57(7), 573–583.
<https://doi.org/10.1641/B570707>
- Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), 1312–1313.
<https://doi.org/10.1093/bioinformatics/btu033>
- Stange, M., Sánchez-Villagra, M. R., Salzburger, W., & Matschiner, M. (2018). Bayesian Divergence-Time Estimation with Genome-Wide Single-Nucleotide Polymorphism Data of Sea Catfishes (Ariidae) Supports Miocene Closure of the Panamanian Isthmus. *Systematic Biology*, 67(4), 681–699. <https://doi.org/10.1093/sysbio/syy006>
- Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M. M. B., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., & Midgley, P. M. (2013). *Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. 109.
- Stuart, Y. E., Losos, J. B., & Algar, A. C. (2012). The island–mainland species turnover relationship. *Proceedings of the Royal Society B: Biological Sciences*, 279(1744), 4071–4077. <https://doi.org/10.1098/rspb.2012.0816>
- Swofford, D. (2017). *PAUP* (4.0) [Computer software].
- Taberlet, P., Fumagalli, L., Wust-Saucy, A., & Cosson, J. (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, 7(4), 453–464.
<https://doi.org/10.1046/j.1365-294x.1998.00289.x>
- Terai, Y., Seehausen, O., Sasaki, T., Takahashi, K., Mizoiri, S., Sugawara, T., Sato, T., Watanabe, M., Konijnendijk, N., Mrosso, H. D. J., Tachida, H., Imai, H., Shichida, Y., & Okada, N. (2006). Divergent Selection on Opsins Drives Incipient Speciation in Lake Victoria Cichlids. *PLoS Biology*, 4(12), e433.
<https://doi.org/10.1371/journal.pbio.0040433>
- Tershy, B. R., Shen, K.-W., Newton, K. M., Holmes, N. D., & Croll, D. A. (2015). The Importance of Islands for the Protection of Biological and Linguistic Diversity. *BioScience*, 65(6), 592–597. <https://doi.org/10.1093/biosci/biv031>
- The 1000 Genomes Project Consortium. (2010). A map of human genome variation from population-scale sequencing. *Nature*, 467(7319), 1061–1073.
<https://doi.org/10.1038/nature09534>
- Toczydlowski, R. H., & Waller, D. M. (2019). Drift happens: Molecular genetic diversity and differentiation among populations of jewelweed (*Impatiens capensis* Meerb.) reflect fragmentation of floodplain forests. *Molecular Ecology*, 28(10), 2459–2475.
<https://doi.org/10.1111/mec.15072>
- Toussaint, R. K., Sage, G. K., Talbot, S. L., & Scheel, D. (2012). Microsatellite marker isolation and development for the giant Pacific Octopus (*Enteroctopus dofleini*). *Conservation Genetics Resources*, 4(3), 545–548. <https://doi.org/10.1007/s12686-011-9588-z>
- Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., & Shipley, P. (2004). micro-checker: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, 4(3), 535–538. <https://doi.org/10.1111/j.1471-8286.2004.00684.x>
- van Zyll de Jong, C. G. (1982). Relationships of amphiberingian shrews of the *Sorex cinereus* group. *Canadian Journal of Zoology*, 60(7), 1580–1587. <https://doi.org/10.1139/z82-208>
- van Zyll de Jong, C. G. (1991). Speciation in the *Sorex cinereus* Group. *The Biology of the Soricidae* (J. S. Findley and T. L. Yates, Eds.). *Museum of Southwestern Biology, Albuquerque, New Mexico*, 65–73.

- Weigelt, P., Jetz, W., & Kreft, H. (2013). Bioclimatic and physical characterization of the world's islands. *Proceedings of the National Academy of Sciences*, 110(38), 15307–15312. <https://doi.org/10.1073/pnas.1306309110>
- Weigelt, Patrick, Steinbauer, M. J., Cabral, J. S., & Kreft, H. (2016). Late Quaternary climate change shapes island biodiversity. *Nature*, 532(7597), 99–102. <https://doi.org/10.1038/nature17443>
- Whitaker, J. (2004). *Sorex cinereus*. *Mammalian Species*, 743, 1–9.
- White, T. A., & Searle, J. B. (2007). Genetic diversity of island populations of the common shrew *Sorex araneus*. *Russian Journal of Theriology*, 6(1), 021–025. <https://doi.org/10.15298/rusjtheriol.06.1.05>
- Whiteley, A. R., Hastings, K., Wenburg, J. K., Frissell, C. A., Martin, J. C., & Allendorf, F. W. (2010). Genetic variation and effective population size in isolated populations of coastal cutthroat trout. *Conservation Genetics*, 11(5), 1929–1943. <https://doi.org/10.1007/s10592-010-0083-y>
- Whittaker, R. J., Triantis, K. A., & Ladle, R. J. (2008). A general dynamic theory of oceanic island biogeography: A general dynamic theory of oceanic island biogeography. *Journal of Biogeography*, 35(6), 977–994. <https://doi.org/10.1111/j.1365-2699.2008.01892.x>
- Winter, D. J. (2012). MMOD: An R library for the calculation of population differentiation statistics. *Molecular Ecology Resources*, 12(6), 1158–1160. <https://doi.org/10.1111/j.1755-0998.2012.03174.x>
- Woolfit, M., & Bromham, L. (2005). Population size and molecular evolution on islands. *Proceedings of the Royal Society B: Biological Sciences*, 272(1578), 2277–2282. <https://doi.org/10.1098/rspb.2005.3217>
- Wright, S. (1931). Evolution in Mendelian Populations. *Genetics*, 16(2), 97–159.
- Yudin. (1969). Taxonomy of Some Species of Shrews (Soricidae) from Palaearctic and Nearctic. *Acta Theriologica*, 14(3), 21–34.
- Zhao, H., Beck, B., Fuller, A., & Peatman, E. (2020). EasyParallel: A GUI platform for parallelization of STRUCTURE and NEWHYBRIDS analyses. *PLOS ONE*, 15(4), e0232110. <https://doi.org/10.1371/journal.pone.0232110>

Appendix A - Rapid allopatric divergence and speciation of an endangered insular shrew (*Sorex pribilofensis*)

Supplementary Figures

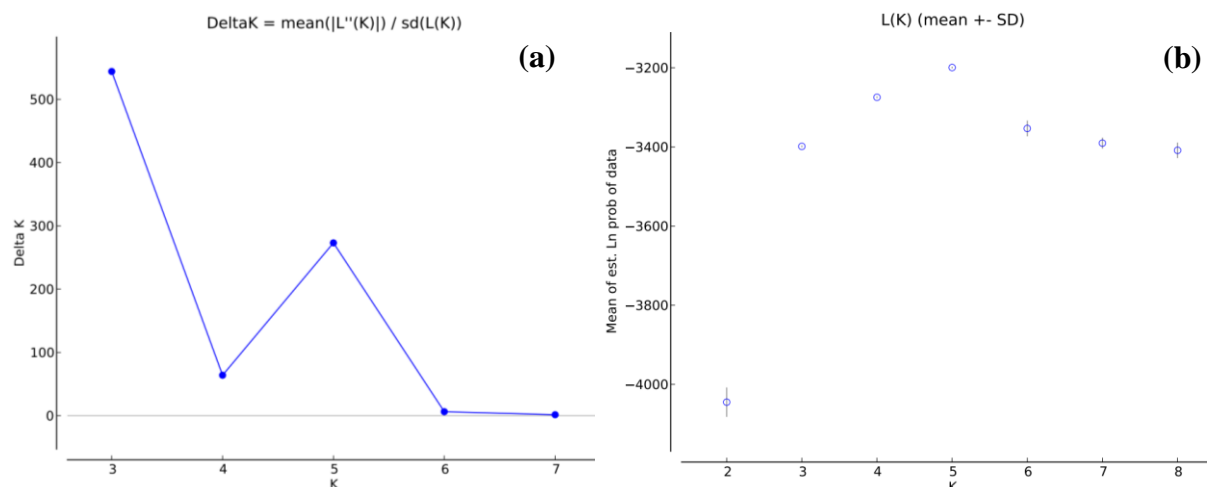


Figure A1. Selection of K for microsatellite Structure analysis

Determination of best value of K for the microsatellite Structure clustering analysis. **(a)** The second order rate of change of the likelihood of each tested value of K (ΔK); based on the Evanno method, ΔK most likely represents the true value of K; **(b)** The mean raw likelihood scores (+/- standard deviation) for each tested value of K. While K=5 has the highest raw likelihood, ΔK is highest for K=3.

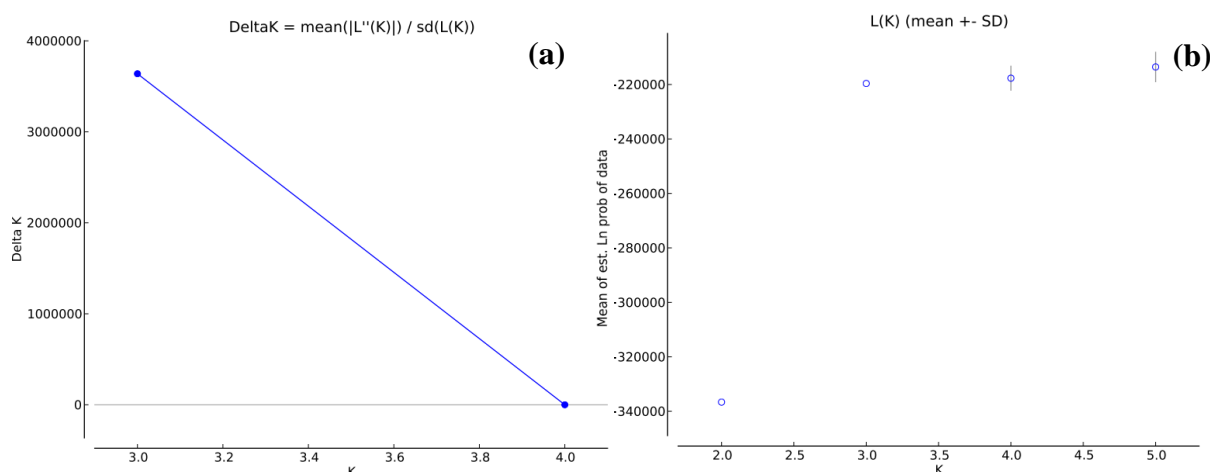


Figure A2. Selection of K for SNP Structure analysis

Determination of best value of K for the SNP Structure clustering analysis. **(a)** The second order rate of change of the likelihood of each tested value of K (ΔK); based on the Evanno method, ΔK most likely represents the true value of K; **(b)** The mean raw likelihood scores (+/- standard deviation) for each tested value of K. While K=5 has the highest raw likelihood, ΔK is highest for K=3.

Supplementary Table

Table A1. Specimens museum numbers and localities

List of specimen museum numbers and localities of each shrew included in this thesis. Inclusion in a specific genetic dataset is indicated by an “X”. For specimens previously sequenced for the Cytb gene, Genbank accession numbers are indicated.

Microsats	CytB	SNPs	Tissue Collection	Tissue Number	Cytb Genbank	Catalog Number	Species	Country	State/Province	Latitude	Longitude	Year
X	X		AF	32746	JN889397	UAM:Mamm:73465	<i>Sorex camtschatica</i>	Russia	Kamchatka Oblast	54.6666667	160.25	1987
X	X		IF	5019	JN889399	UAM:Mamm:84134	<i>Sorex camtschatica</i>	Russia	Magadan Oblast	59.7166667	150.85	2002
	X		AF	6569	AY014919		<i>Sorex camtschatica</i>					
X			AF	6596		UAM:Mamm:29148	<i>Sorex camtschatica</i>	Russia	Magadan Oblast	59.7333333	150.8666667	1994
	X		AF	6631	AY014917	UAM:Mamm:29149	<i>Sorex camtschatica</i>	Russia	Magadan Oblast	59.7166667	150.8666667	1994
X	X		AF	6632	AY014918	UAM:Mamm:29150	<i>Sorex camtschatica</i>	Russia	Magadan Oblast	59.7333333	150.8666667	1994
X	X		AF	41349	JN889398	UAM:Mamm:80338	<i>Sorex camtschatica</i>	Russia	Magadan Oblast	60.7566667	151.7838889	2000
	X		CSW	4664	AY014916	UWBM:Mamm:39234	<i>Sorex camtschatica</i>	Russia	Kamchatka Krai	53.25	158.183	2011
	X	X	NK	123278	JN889593	MSB:Mamm:143101	<i>Sorex cinereus</i>	United States	Alaska	66.35	-150.46	2005
X		X	NK	196372		MSB:Mamm:223704	<i>Sorex cinereus</i>	United States	Alaska	67.7039	-156.9665	2010
X	X	X	NK	196469		MSB:Mamm:223732	<i>Sorex cinereus</i>	United States	Alaska	68.2119	-159.8689	2010
X	X	X	NK	196499		MSB:Mamm:223718	<i>Sorex cinereus</i>	United States	Alaska	68.2047	-159.8609	2010
X	X	X	NK	196562		MSB:Mamm:222039	<i>Sorex cinereus</i>	United States	Alaska	65.3784	-163.2283	2010
X	X	X	NK	213804	JQ778872	MSB:Mamm:248416	<i>Sorex cinereus</i>	United States	Alaska	54.85374	-163.41527	2011
X	X	X	NK	211848		MSB:Mamm:278698	<i>Sorex cinereus</i>	United States	Alaska	66.36026	-150.44716	2013
X	X	X	NK	211849		MSB:Mamm:278699	<i>Sorex cinereus</i>	United States	Alaska	66.36026	-150.44716	2013
X	X	X	NK	211851		MSB:Mamm:278701	<i>Sorex cinereus</i>	United States	Alaska	66.36026	-150.44716	2013
X	X	X	NK	211852		MSB:Mamm:278702	<i>Sorex cinereus</i>	United States	Alaska	66.36026	-150.44716	2013
X	X	X	NK	211862		MSB:Mamm:278712	<i>Sorex cinereus</i>	United States	Alaska	66.36026	-150.44716	2013
X	X	X	NK	211866		MSB:Mamm:278716	<i>Sorex cinereus</i>	United States	Alaska	66.36026	-150.44716	2013
X	X	X	NK	211873		MSB:Mamm:278723	<i>Sorex cinereus</i>	United States	Alaska	66.36026	-150.44716	2013

X	X	X	NK	211874		MSB:Mamm:278724	<i>Sorex cinereus</i>	United States	Alaska	66.36026	-150.44716	2013
X	X	X	NK	211876		MSB:Mamm:278726	<i>Sorex cinereus</i>	United States	Alaska	66.36026	-150.44716	2013
X			UAM	7777		UAM:Mamm:7777	<i>Sorex jacksoni</i>	United States	Alaska	63.295	-168.6922222	1954
X			UAM	7792		UAM:Mamm:7792	<i>Sorex jacksoni</i>	United States	Alaska	63.295	-168.6922222	1959
X			UAM	7794		UAM:Mamm:7794	<i>Sorex jacksoni</i>	United States	Alaska	63.295	-168.6922222	1959
X	X		UAM	21508	JN889520	UAM:Mamm:21508	<i>Sorex jacksoni</i>	United States	Alaska	63.43333333	-171.1	1959
X			UAM	21509		UAM:Mamm:21509	<i>Sorex jacksoni</i>	United States	Alaska	63.43333333	-171.1	1959
	X		UAM	21510	JN889521	UAM:Mamm:21510	<i>Sorex jacksoni</i>	United States	Alaska	63.43333333	-171.1	1959
	X		UAM	21518	AY014922	UAM:Mamm:21518	<i>Sorex jacksoni</i>	United States	Alaska	63.78333333	-171.75	1960
	X		UAM	21520	AY014923	UAM:Mamm:21520	<i>Sorex jacksoni</i>	United States	Alaska	63.78333333	-171.75	1960
	X		UAM	21521	AY014924	UAM:Mamm:21521	<i>Sorex jacksoni</i>	United States	Alaska	63.78333333	-171.75	1961
	X		UAM	21522	AY014925	UAM:Mamm:21522	<i>Sorex jacksoni</i>	United States	Alaska	63.78333333	-171.75	1961
X	X		UAM	21526	AY014926	UAM:Mamm:21526	<i>Sorex jacksoni</i>	United States	Alaska	63.775	-171.6930556	1952
X			IF	5169		UAM:Mamm:84317	<i>Sorex portenkoi</i>	Russia	Chukotka Autonomous Okrug	64.80555556	177.5541667	2002
	X	X	IF	5187	JN889541	UAM:Mamm:84327	<i>Sorex portenkoi</i>	Russia	Chukotka Autonomous Okrug	64.80555556	177.5541667	2002
X	X		IF	5222	JN889538	UAM:Mamm:84282	<i>Sorex portenkoi</i>	Russia	Chukotka Autonomous Okrug	64.81638889	177.5455556	2002
X	X		IF	5228	JN889539	UAM:Mamm:84286	<i>Sorex portenkoi</i>	Russia	Chukotka Autonomous Okrug	64.81638889	177.5455556	2002
X	X	X	IF	5243	JN889540	UAM:Mamm:84290	<i>Sorex portenkoi</i>	Russia	Chukotka Autonomous Okrug	64.81638889	177.5455556	2002
X			IF	7555		UAM:Mamm:83822	<i>Sorex portenkoi</i>	Russia	Chukotka Autonomous Okrug	64.41666667	-172.5333333	2002
X			IF	7562		UAM:Mamm:83829	<i>Sorex portenkoi</i>	Russia	Chukotka Autonomous Okrug	64.41666667	-172.5333333	2002
X	X	X	IF	7570	JN889537	UAM:Mamm:83837	<i>Sorex portenkoi</i>	Russia	Chukotka Autonomous Okrug	64.41666667	-172.5333333	2002
X	X		IF	7573	JN889652	UAM:Mamm:83840	<i>Sorex portenkoi</i>	Russia	Chukotka Autonomous Okrug	64.41666667	-172.5333333	2002

X			IF	7574		UAM:Mamm:83841	<i>Sorex portenkoi</i>	Russia	Chukotka Autonomous Okrug	64.41666667	-172.5333333	2002
	X	X	IF	7582		UAM:Mamm:83849	<i>Sorex portenkoi</i>	Russia	Chukotka Autonomous Okrug	64.41666667	-172.5333333	2002
	X	X	IF	7583		UAM:Mamm:83850	<i>Sorex portenkoi</i>	Russia	Chukotka Autonomous Okrug	64.41666667	-172.5333333	2002
X	X	X	IF	7584	JN889653	UAM:Mamm:83851	<i>Sorex portenkoi</i>	Russia	Chukotka Autonomous Okrug	64.41666667	-172.5333333	2002
X	X	X	IF	7585	JN889701	UAM:Mamm:83852	<i>Sorex portenkoi</i>	Russia	Chukotka Autonomous Okrug	64.41666667	-172.5333333	2002
X	X	X	IF	7590		UAM:Mamm:83857	<i>Sorex portenkoi</i>	Russia	Chukotka Autonomous Okrug	64.41666667	-172.5333333	2002
X			IF	7591		UAM:Mamm:83858	<i>Sorex portenkoi</i>	Russia	Chukotka Autonomous Okrug	64.41666667	-172.5333333	2002
	X	X	IF	7597		UAM:Mamm:83864	<i>Sorex portenkoi</i>	Russia	Chukotka Autonomous Okrug	64.41666667	-172.5333333	2002
	X	X	IF	7598		UAM:Mamm:83867	<i>Sorex portenkoi</i>	Russia	Chukotka Autonomous Okrug	64.41666667	-172.5333333	2002
X			UAM	22591		UAM:Mamm:22591	<i>Sorex pribilofensis</i>	United States	Alaska	57.16666667	-170.25	1986
X	X		UAM	22593	AY014931	UAM:Mamm:22593	<i>Sorex pribilofensis</i>	United States	Alaska	57.16666667	-170.25	1986
X	X		UAM	22594	AY014932	UAM:Mamm:22594	<i>Sorex pribilofensis</i>	United States	Alaska	57.16666667	-170.25	1986
X			UAM	50348		UAM:Mamm:50348	<i>Sorex pribilofensis</i>	United States	Alaska	57.24722222	-170.0972222	1987
X	X	X	NK	234251		MSB:Mamm:302936	<i>Sorex pribilofensis</i>	United States	Alaska	57.12	-170.28	2014
X	X	X	NK	234252		MSB:Mamm:305188	<i>Sorex pribilofensis</i>	United States	Alaska	57.15	-170.34	2014
X	X	X	NK	234253		MSB:Mamm:305189	<i>Sorex pribilofensis</i>	United States	Alaska	57.12	-170.28	2014
X	X	X	NK	234254		MSB:Mamm:302937	<i>Sorex pribilofensis</i>	United States	Alaska	57.138	-170.293	2014
X			NK	234255		MSB:Mamm:305190	<i>Sorex pribilofensis</i>	United States	Alaska	57.15	-170.255	2014
X	X	X	NK	234256		MSB:Mamm:305191	<i>Sorex pribilofensis</i>	United States	Alaska	57.12	-170.28	2014
X		X	NK	234351		MSB:Mamm:302942	<i>Sorex pribilofensis</i>	United States	Alaska	57.12	-170.28	2014
X			NK	234352		MSB:Mamm:302943	<i>Sorex pribilofensis</i>	United States	Alaska	57.12	-170.28	2014
X	X	X	NK	234353		MSB:Mamm:305195	<i>Sorex pribilofensis</i>	United States	Alaska	57.12	-170.28	2014

X	X	X	NK	234354	MSB:Mamm:302944	<i>Sorex pribilofensis</i>	United States	Alaska	57.12	-170.28	2014
X			NK	234355	MSB:Mamm:305196	<i>Sorex pribilofensis</i>	United States	Alaska	57.12	-170.28	2015
X			NK	234356	MSB:Mamm:302945	<i>Sorex pribilofensis</i>	United States	Alaska	57.12	-170.28	2014
X	X	X	NK	234357	MSB:Mamm:305197	<i>Sorex pribilofensis</i>	United States	Alaska	57.12	-170.28	2014
X	X	X	NK	234395	MSB:Mamm:302947	<i>Sorex pribilofensis</i>	United States	Alaska	57.12	-170.28	2015
X	X	X	NK	234400	MSB:Mamm:302946	<i>Sorex pribilofensis</i>	United States	Alaska	57.12	-170.28	2015
X			NK	234401	MSB:Mamm:305762	<i>Sorex pribilofensis</i>	United States	Alaska	57.12	-170.28	Not recorded 2015
X	X	X	NK	234407	MSB:Mamm:305764	<i>Sorex pribilofensis</i>	United States	Alaska	57.12	-170.28	
X		X	NK	234408	MSB:Mamm:305765	<i>Sorex pribilofensis</i>	United States	Alaska	57.12	-170.28	2015
	X	X	NK	305141	MSB:Mamm:331411	<i>Sorex pribilofensis</i>	United States	Alaska	57.12415	-170.28328	2019
	X	X	NK	305142	MSB:Mamm:331432	<i>Sorex pribilofensis</i>	United States	Alaska	57.11989	-170.27794	2019
	X	X	NK	305143	MSB:Mamm:331475	<i>Sorex pribilofensis</i>	United States	Alaska	57.15176	-170.22151	2019
	X	X	NK	305144	MSB:Mamm:331478	<i>Sorex pribilofensis</i>	United States	Alaska	57.1516	-170.22205	2019
	X	X	NK	305145	MSB:Mamm:332151	<i>Sorex pribilofensis</i>	United States	Alaska	57.21195	-170.15201	2019
	X	X	NK	305146	MSB:Mamm:331496	<i>Sorex pribilofensis</i>	United States	Alaska	57.21195	-170.15216	2019
	X	X	NK	305147	MSB:Mamm:331390	<i>Sorex pribilofensis</i>	United States	Alaska	57.21202	-170.15222	2019
	X	X	NK	305148	MSB:Mamm:331399	<i>Sorex pribilofensis</i>	United States	Alaska	57.21203	-170.15225	2019
	X	X	NK	305149	MSB:Mamm:331416	<i>Sorex pribilofensis</i>	United States	Alaska	57.21204	-170.15222	2019
	X	X	NK	305150	MSB:Mamm:331421	<i>Sorex pribilofensis</i>	United States	Alaska	57.21179	-170.15166	2019
	X	X	NK	305151	MSB:Mamm:331424	<i>Sorex pribilofensis</i>	United States	Alaska	57.21195	-170.15216	2019
	X	X	NK	305152	MSB:Mamm:332412	<i>Sorex pribilofensis</i>	United States	Alaska	57.21196	-170.15213	2019
	X	X	NK	305153	MSB:Mamm:331445	<i>Sorex pribilofensis</i>	United States	Alaska	57.21195	-170.15202	2019
	X	X	NK	305154	MSB:Mamm:331446	<i>Sorex pribilofensis</i>	United States	Alaska	57.15173	-170.22145	2019
	X	X	NK	305155	MSB:Mamm:331447	<i>Sorex pribilofensis</i>	United States	Alaska	57.21196	-170.15213	2019
		X	NK	305156	MSB:Mamm:331448	<i>Sorex pribilofensis</i>	United States	Alaska	57.21191	-170.15205	2019
	X	X	NK	305157	MSB:Mamm:331449	<i>Sorex pribilofensis</i>	United States	Alaska	57.21182	-170.1517	2019
	X	X	NK	305158	MSB:Mamm:332423	<i>Sorex pribilofensis</i>	United States	Alaska	57.21202	-170.15225	2019

	X	X	NK	305159		MSB:Mamm:332424	<i>Sorex pribilofensis</i>	United States	Alaska	57.21178	-170.15167	2019
	X	X	NK	305160		MSB:Mamm:331450	<i>Sorex pribilofensis</i>	United States	Alaska	57.15174	-170.22147	2019
	X	X	NK	305161		MSB:Mamm:331451	<i>Sorex pribilofensis</i>	United States	Alaska	57.15171	-170.22157	2019
	X	X	NK	305162		MSB:Mamm:331452	<i>Sorex pribilofensis</i>	United States	Alaska	57.21162	-170.15146	2019
	X	X	NK	305163		MSB:Mamm:331453	<i>Sorex pribilofensis</i>	United States	Alaska	57.21201	-170.15222	2019
	X	X	NK	305164		MSB:Mamm:331454	<i>Sorex pribilofensis</i>	United States	Alaska	57.15157	-170.222	2019
	X	X	NK	305165		MSB:Mamm:331913	<i>Sorex pribilofensis</i>	United States	Alaska	57.15157	-170.222	2019
	X	X	NK	305166		MSB:Mamm:331916	<i>Sorex pribilofensis</i>	United States	Alaska	57.15157	-170.222	2019
	X	X	NK	305167		MSB:Mamm:331917	<i>Sorex pribilofensis</i>	United States	Alaska	57.21202	-170.15225	2019
	X	X	NK	305182		MSB:Mamm:331460	<i>Sorex pribilofensis</i>	United States	Alaska	57.12	-170.2781	2018
	X	X	NK	305183		MSB:Mamm:331461	<i>Sorex pribilofensis</i>	United States	Alaska	57.12	-170.2781	2018
	X	X	NK	305184		MSB:Mamm:331462	<i>Sorex pribilofensis</i>	United States	Alaska	57.12	-170.2781	2018
	X	X	NK	305186		MSB:Mamm:331349	<i>Sorex pribilofensis</i>	United States	Alaska	57.12	-170.2781	2018
X	X	X	NK	196254		MSB:Mamm:221848	<i>Sorex ugyunak</i>	United States	Alaska	68.1174	-154.124	2010
X	X	X	NK	196256		MSB:Mamm:221849	<i>Sorex ugyunak</i>	United States	Alaska	68.1174	-154.124	2010
X	X	X	NK	196262		MSB:Mamm:221845	<i>Sorex ugyunak</i>	United States	Alaska	68.1174	-154.124	2010
X	X	X	NK	196265		MSB:Mamm:221839	<i>Sorex ugyunak</i>	United States	Alaska	68.1174	-154.124	2010
X	X	X	NK	196266		MSB:Mamm:221842	<i>Sorex ugyunak</i>	United States	Alaska	68.1174	-154.124	2010
X	X	X	NK	196267		MSB:Mamm:221843	<i>Sorex ugyunak</i>	United States	Alaska	68.1174	-154.124	2010
X	X	X	NK	196282		MSB:Mamm:221794	<i>Sorex ugyunak</i>	United States	Alaska	68.1174	-154.124	2010
X	X	X	NK	196283	JQ778933	MSB:Mamm:221828	<i>Sorex ugyunak</i>	United States	Alaska	68.1174	-154.124	2010
X	X	X	NK	196285	JQ778935	MSB:Mamm:221819	<i>Sorex ugyunak</i>	United States	Alaska	68.1174	-154.124	2010
X	X	X	NK	196301		MSB:Mamm:223685	<i>Sorex ugyunak</i>	United States	Alaska	68.1174	-154.124	2010
X	X	X	NK	196394		MSB:Mamm:223717	<i>Sorex ugyunak</i>	United States	Alaska	67.7188	-156.1409	2010
X	X	X	NK	196477		MSB:Mamm:223768	<i>Sorex ugyunak</i>	United States	Alaska	68.2047	-159.8609	2010
X	X	X	NK	196486		MSB:Mamm:223706	<i>Sorex ugyunak</i>	United States	Alaska	68.2119	-159.8689	2010
X	X	X	NK	196564		MSB:Mamm:222005	<i>Sorex ugyunak</i>	United States	Alaska	65.3784	-163.2283	2010

X	X	X	NK	196605		MSB:Mamm:223122	<i>Sorex ugyunak</i>	United States	Alaska	65.3784	-163.2283	2010
X	X	X	NK	196621		MSB:Mamm:223124	<i>Sorex ugyunak</i>	United States	Alaska	65.3784	-163.2283	2010
X	X	X	NK	196637		MSB:Mamm:223120	<i>Sorex ugyunak</i>	United States	Alaska	65.3784	-163.2283	2010
X	X	X	NK	123279	JF430996	MSB:Mamm:143102	<i>Sorex ugyunak</i>	United States	Alaska	66.35	-150.46	2005
X	X	X	NK	213919	JQ778950	MSB:Mamm:247467	<i>Sorex ugyunak</i>	United States	Alaska	70.721503	-153.836542	2011
		X	NK	234248		MSB:Mamm:305187	<i>Sorex ugyunak</i>	United States	Alaska	71.25377	-156.61072	Not recorded
	X	X	NK	234250		MSB:Mamm:302935	<i>Sorex ugyunak</i>	United States	Alaska	71.25378	-156.61069	2013

Appendix B - Genetic drift drives differentiation of an endangered insular shrew (*Sorex pribilofensis*)

Supplementary Table

Table B1. Significance and goodness of fit of linear regressions

Significance and goodness of fit of linear regressions between genetic variability and genetic differentiation for the microsatellite and neutral SNP datasets.

Microsatellites					
y	x	Multiple R-squared	F-statistic	p-value	Adjusted p-value
Jost's D	H _o	0.9055	28.76	0.01269	0.01269
Nei's F _{ST}	H _o	0.9648	82.19	0.002835	0.003189
Nei's G _{ST}	H _o	0.9719	103.7	0.002019	0.003189
Jost's D	H _e	0.972	104	0.002008	0.003189
Nei's F _{ST}	H _e	0.9769	127	0.001498	0.003189
Nei's G _{ST}	H _e	0.9697	96.04	0.002258	0.003189
Jost's D	Ar	0.9674	88.9	0.002528	0.003189
Nei's F _{ST}	Ar	0.9905	313.5	0.000393	0.003189
Nei's G _{ST}	Ar	0.9852	200.2	0.000765	0.003189
Neutral SNPs					
y	x	Multiple R-squared	F-statistic	p-value	Adjusted p-value
Jost's D	H _o	0.9619	25.24	0.1251	0.1251
Nei's F _{ST}	H _o	0.9994	1558	0.01612	0.04837
Nei's G _{ST}	H _o	0.9997	3554	0.01068	0.04805
Jost's D	H _e	0.9729	35.96	0.1052	0.1184
Nei's F _{ST}	H _e	0.9999	28750	0.003754	0.03379
Nei's G _{ST}	H _e	0.9977	433.4	0.03056	0.06252
Jost's D	Ar	0.9865	72.91	0.07422	0.09542
Nei's F _{ST}	Ar	0.997	335.2	0.03474	0.06252
Nei's G _{ST}	Ar	0.9907	106.4	0.06154	0.09231